Molecular analysis of coenzyme A ligase from benzoate-metabolizing *Sorbus aucuparia* cell cultures

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat)
genehmigte
D i s s e r t a t i o n

von Hussein Ramadan
aus Rabta, Libyen

1. Referee: Prof. Dr. Ludger Beerhues
2. Referee: apl. Prof. Dr. Dirk Selmar

eingereicht am: 14.08.2006
mündliche Prüfung (Disputation) am: 19.10.2006
Druckjahr : 2006

Parts of this work have previously been published with permission of the Faculty of Life Science, represented by the mentor of this work:

Presentations

-Short lecture

Ramadan, H , Beerhues, L (2005)
Coenzyme A ligases involved in benzoic acid biosynthesis
XVII International Botanical Congress
Austria Center Vienna, 17 - 23 July 2005

-Poster

Ramadan, H , Beerhues, L (2004)
Coenzyme A ligases involved in benzoic acid biosynthesis
Botanikertagung der Deutschen Botanischen Gesellschaft und der Vereinigung für Angewandte Botanik
Braunschweig, 05-10 September 2004

ACKNOWLEDGEMENTS

I would like to express my gratitude to many individuals who assisted me throughout the course of my doctoral research. First and foremost, I would like to express my gratitude to my advisor, Professor Dr. Ludger Beerhues, for the wonderful opportunity of participating in this exciting research endeavour and for constant guidance, unwavering moral support during my years as PhD student. More thanks for his patience to correct from the beginning to the end my thesis tirelessly, and for the extraordinary advice, caring, encouragement, and affection he bestowed on me.

My sincere thanks go to Prof. Dr. Benye Liu and Dr. Heiko Schwarz for invaluable assistance, suggestions, and discussion. Especially, for the first success in molecular biology of my career.

Sincere gratitude is also extended to Dr. Rainer Lindigkeit for special attention, personal help, using his time and effort to show me different computer programs.

Many thanks go to all staff and faculty members of the Institute of Pharmaceutical Biology of TU Braunschweig, present or past, who in one way or another have contributed a lot for my life in the institute to be so enjoyable and unforgettable, and making this study a successful experience, specially I extend immense appreciation and thanks to Mrs. Ines Rahaus for providing a stimulating and fun environment and for all her help.

I also owe heartfelt exceptional thanks to my wife, Dr. Zakia Boubakir, for her patient love. Without her support and love, I would have not been able to complete this work. I also thank my kids Retaj and Roba for their patience and emotional support.

My sincere thanks are to my family: my parents for educating me with aspects from both arts and sciences and for unconditional support and encouragement to finish my PhD work; my brothers and my sisters for their ever-loving care towards me throughout.

Lastly, I would like to thank the Libyan Government for providing financial support for my doctoral study and the Deutsche Forschungsgemeinschaft (DFG) for funding this research project.

List of Tables

List of Figures

ABBREVIATIONS

λ	Wavelength
µ	Micro
µM	Micromolar
µkat	Microkatals
ACS	Acridone synthase
AMP	Adenosine monophosphate
APS	Ammonium Peroxy Disulphate
ATP	Adenosine triphosphate
BBS	Bibenzyl synthase
BIS	Biphenyl synthase
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BPS	Benzophenone synthase
BSA	Bovine Serum Albumin
°C	Degree Celsius
cDNA	Complementary DNA
4CL	4-coumarate:coenzyme A ligase
C4H	Cinnamate 4-hydroxylase
CHS	Chalcone synthase
CIP	Calf Intestinal Phosphatase
CKs	Cytokinins
CoA	Coenzyme A
Da	Dalton
DGP	Degenerate Primers
DMSO	Dimethyl sulfoxide
DTT	1,4-Dithiothreitol
DNA	Deoxyribo nucleic acid
λDNA	Lambda deoxyribo nucleic acid
dNTP	Deoxynucleoside triphosphate
ddNTPs	Dye-labeled chain-terminating nucleotides

DOP-PCR	Degenerate Oligonucleotide Primed PCR
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast Protein Liquid Chromatography
GBA	3-geranyl-4-hydroxybenzoate
GPP	Geranyl diphosphate
GSP	Gene Specific Primers
GTC	Guanidinium thiocyanate
h	Hour
4HB	4-hydroxybenzoate
HIV	Human immunodeficiency virus
HMGR	3-hydroxy-3-methylglutaryl-coenzyme A reductase
HPLC	High performance liquid chromatography
IPTG	Isopropyl-ß-D-thiogalactopyranoside
kb	Kilo base pairs
K_{cat}	Catalytic constant
K_{cat}/K_m	Ration Catalytic constant / kinetic constant
kDa	Kilodalton
Km	Kinetic constant Michaelis-Konstante
LB	Luria-Bertani
LS	Linsmeier und Skoog
M	Molar
MCS	Multiple cloning site
MeSA	Methyl salicylate
min	Minute
mM	Millimolar
M_r	Relative molecular mass
mRNA	messenger RNA
NADPH	Nicotinamide Adenine Dinucleotide Phosphate Hydrogen
NCPI	National Center of Biological Information
OD	Optical density
ORF	Open Reading Frame

PAGE	Polyacrylamide gel electrophoresis
PAD	Photo diode array detection
PAL	L-phenylalanine ammonia–lyase
PCR	Polymerase Chain Reaction
Pfu DNA polymerase	*Pyrrococcus furiosus* polymerase
pkat	picokatal
PKS	Polyketide synthase
PRSA	Penicillin-resistant *Staphylococcus aureus*
PSSA	Penicillin-susceptible *S. aureus*
RACE	Rapid Amplification of cDNA Ends
$RACE_{long+short}$	RACE anchor primer
RLM-RACE	RNA Ligase mediated rapid amplification of cDNA Ends
RNA	Ribonucleic acid
RNAse	Ribonuclase inhibitor
rpm	Revolutions per minute
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
R_t value	HPLC retention time
SA	Salicylic acid (2-hydroxybenzoic acid)
SAG	β-O-D-glucosylsalicylic acid
SAR	Systemic Acquired Resistance
SBP	Substrate binding pocket
SDS	Sodium dodecyl sulphate
SMART-RACE	Switching Mechanism At 5' end of RNA Template
STS	Stilbene synthase
TAE	Tris-Acetate-EDTA
Taq DNA polymerase	Thermus aquaticus polymerase
TAP	Tobacco Acid Pyrophosphatase
TBE	Tris-Borate-EDTA
***t*-CA**	*Trans*-cinnamic acid
TdT	Terminal Deoxynucleotidyl Transferase
TEMED	*N,N,N',N'*-Tetramethylethylendiamin
***T*m**	Melting temperature

Tris	Tris-(hydoxymethyl)aminomethan
tRNA	Total RNA
V_o	Void volume
V_e	Elution volume
V_{max}	Maximum velocity
v/v	Volume per volume
UV	Ultraviolett
X-gal	5-bromo-4-chloro-3-indolyl-ß-D-galactoside

Amino acids:

Abbreviation	**Amino acid**
A = Ala	Alanin
C = Cys	Cystein
D = Asp	Asparaginic acid
E = Glu	Glutaminic acid
F = Phe	Phenylalanin
G = Gly	Glycin
H = His	Histidin
I = Ile	Isoleucin
K = Lys	Lysin
L = Leu	Leucin
M = Met	Methionin
N = Asn	Asparagin
P = Pro	Prolin
Q = Gln	Glutamin
R = Arg	Arginin
S = Ser	Serin
T = Thr	Threonin
V = Val	Valin
W = Trp	Tryptophan
Y = Tyr	Tyrosin

1. INTRODUCTION

1.1 Secondary Metabolism

Metabolism is a chemical aspect of life that produces metabolites and energy by chemical reactions catalyzed by enzymes. Metabolic reactions that synthesize basic metabolites such as nucleic acids, carbohydrates, proteins, and lipids are shared by all organisms and are summarized as primary metabolism. Those that synthesize secondary metabolites such as antibiotics, toxins, pigments, and hormones are specific to individual organisms. The term secondary metabolite usually refers to a biological molecule that is not involved in primary metabolism as the basic machinery of life. Although secondary metabolites often represent only a small fraction of the total biomass of an organism, they are at the basis of ecological specialization because they can affect species distribution and feeding patterns, as well as reproductive success and community organization. Secondary metabolites have biological activities that are valuable for medical and agricultural uses and find application in chemical industries.

In plants, many of the organic compounds formed are not required for the primary physiological functions supporting life such as photosynthesis, respiration, growth, and reproduction. These non-essential compounds are the secondary metabolites and associated with the plant's adaptation to its environment. They benefit the plant as chemicals toxic to predators or pigments attractive to pollinators and may ultimately determine the viability and fitness of the plant.

Plant secondary metabolites, such as phenylpropanoids, terpenoids, and alkaloids, are classified according to the type of molecules from which they are derived (Croteau et al., 2000). Their diversity results in part from molecular modifications (oxidation, acylation, glucosylation, and methylation) of functional groups of their precursors that are either primary metabolites or intermediate secondary compounds (De Luca and St. Pierre, 2000). Many secondary metabolic pathways do not occur universally among plant species but are restricted to a few taxa. A metabolic pathway is reconstructed by predicting intermediate reactions stepwise in the direction reverse to biosynthesis, from a given secondary compound to a starting basic metabolite. The present work focuses on benzoic acids as phenylpropanoid derivatives.

1.2 Phenylpropanoid Metabolism

Phenolic compounds produced via the pathway initiated by L-phenylalanine ammonia–lyase (PAL) are referred to as phenylpropanoid compounds. Phenylpropanoids include coumarins, flavonoids, stilbenes, and lignin. Compounds belonging to these groups incorporate all nine carbon atoms from cinnamic acid into their respective skeletons. The first committed step in the biosynthesis of phenylpropanoid compounds is the conversion of L–phenylalanine to *trans*-cinnamic acid by L-phenylalanine ammonia–lyase (Fig.1-1). PAL controls the flux of carbon into the pathway and therefore the efficiency of the whole pathway (Hahlbrock and Scheel, 1989). The second step in the phenylpropanoid pathway, the hydroxylation of *trans*-cinnamic acid to 4-coumaric acid, is catalyzed by a cytochrome P450 monooxygenase, cinnamic acid 4-hydroxylase (Russell and Conn, 1967; Fahrendorf and Dixon, 1993; Teutsch et al., 1993). The enzyme 4-coumarate:coenzyme A ligase (4CL) catalyzes the last reaction of the general phenylpropanoid pathway in plants. 4CL catalyzes the conversion of 4-coumarate (4-hydroxycinnamate) and other substituted cinnamates, such as caffeate (3,4 dihydroxycinnamate) and ferulate (3-methoxy-4-hydroxycinnamate), into the corresponding CoA thiol esters. 4CL-catalyzed CoA ester formation takes place via a two-step reaction. During the first step, 4-coumarate and ATP form a coumaroyl-adenylate intermediate with the simultaneous release of pyrophosphate. In the second step, the coumaroyl group is transferred to the sulfhydryl group of CoA, and AMP is released (Becker et al., 1991; Knobloch, et al., 1975). The mechanism of formation of an adenylate intermediate is common among a number of enzymes with divergent functions, including luciferases, fatty acyl-CoA ligases, acetyl-CoA ligases, and specialized domains within peptide synthetase multienzymes. Despite their low overall amino acid sequence identity, similar reaction mechanisms and the presence of conserved peptide motifs were used as criteria to classify them in a superfamily of adenylate-forming enzymes (Fulda, et al., 1994). 4-Coumarate:coenzyme A ligase reveals two large, distinct classes, class I and class II. Class I in 4CLs *Arabidopsis*, poplar, and soybean are more associated with biosynthesis of lignins and other phenylpropanoids whereas class II 4CLs have been generally associated with flavonoid biosynthesis.

CoA thiol esters of 4-coumarate and other hydroxycinnamates formed by 4CL serve as substrates for entry into various branch pathways. For example, hydroxycinnamoyl CoA esters enter the monolignol biosynthetic pathway by undergoing two successive reductive steps. The first one is carried out by hydroxycinnamoyl-CoA reductase which catalyzes the NADPH-dependent reduction of hydroxycinnamoyl CoA esters to hydroxycinnamaldehydes.

These products are taken by hydroxycinnamyl alcohol dehydrogenase to produce the hydroxycinnamyl alcohols or monolignols (4-coumaryl, coniferyl, and sinapyl alcohols). Lignin is integrated into the plant secondary cell wall where it provides structural rigidity to plant tissues and enables tracheary elements to withstand the tension generated during transpiration.

From 4-coumaryoyl-CoA, the pathway also branches to the flavonoid and stilbene biosynthetic pathways. The entry enzyme into the flavonoid pathway is chalcone synthase (CHS). The formation of the C_{15} skeleton resulting from the combination of 4-coumaroyl-CoA (C_9) with three molecules of decarboxylated malonyl-CoA (3 x C2) is catalyzed by chalcone synthase (CHS). Thus, all flavonoid derivatives contain the phenylpropanoid nucleus (B-ring) from phenylalanine and a second aromatic ring (A-ring) derived from the three molecules of malonyl-CoA. After chalcone isomerase has formed a flavanone from the CHS product, further modifications lead to various flavonoid subgroups such as isoflavonoids which are produced by isoflavone synthase (Goodwin and Mercer, 1983). Chalcone synthase can commonly use both 4-coumaroyl-CoA and cinnamoyl-CoA, giving rise to naringenin chalcone or pinocembrin chalcone, respectively (Fliegmann et al., 1992; Kodan et al., 2002). Chalcone synthase is encoded in plants either as a single copy gene (*Arabidopsis thaliana*; Koch et al., 2001), multiple gene copies (*Petunia hybrida*; Koes et al., 1989) or gene family (leguminous plants; Akada et al., 1995).

Although phenylpropanoid derivatives are not usually volatile, some of them with a reduced carboxyl group at C-9 (e.g. aldehydes, alcohols) or those containing alkyl additions to the hydroxyl groups of the phenyl ring or to the carboxyl group (i.e. ethers, esters) are volatile (Dudareva and Pichersky, 2000). Methylisoeugenol, methylchavicol and methylcinnamate are frequently found and investigated in floral scents of *Clarkia breweri* and *Ocimum basilicum*, together with benzenoid esters, terpenoids, fatty acid derivatives, C_5-branched compounds, and various nitrogen and sulfur containing compounds (Dudareva et al., 2004).

Sinapoylesters (sinapoylmalate, sinapoylcholine) have been found predominantly in *Brassicaceae*, where these esters are among the major soluble phenylpropanoid metabolites to be accumulated (Lorenzen et al., 1996). While the function of sinapoylcholine as such is not yet elucidated, sinapoylmalate has been found to act as an UV-B protectant in vegetative plant tissues (Booij et al., 2000).

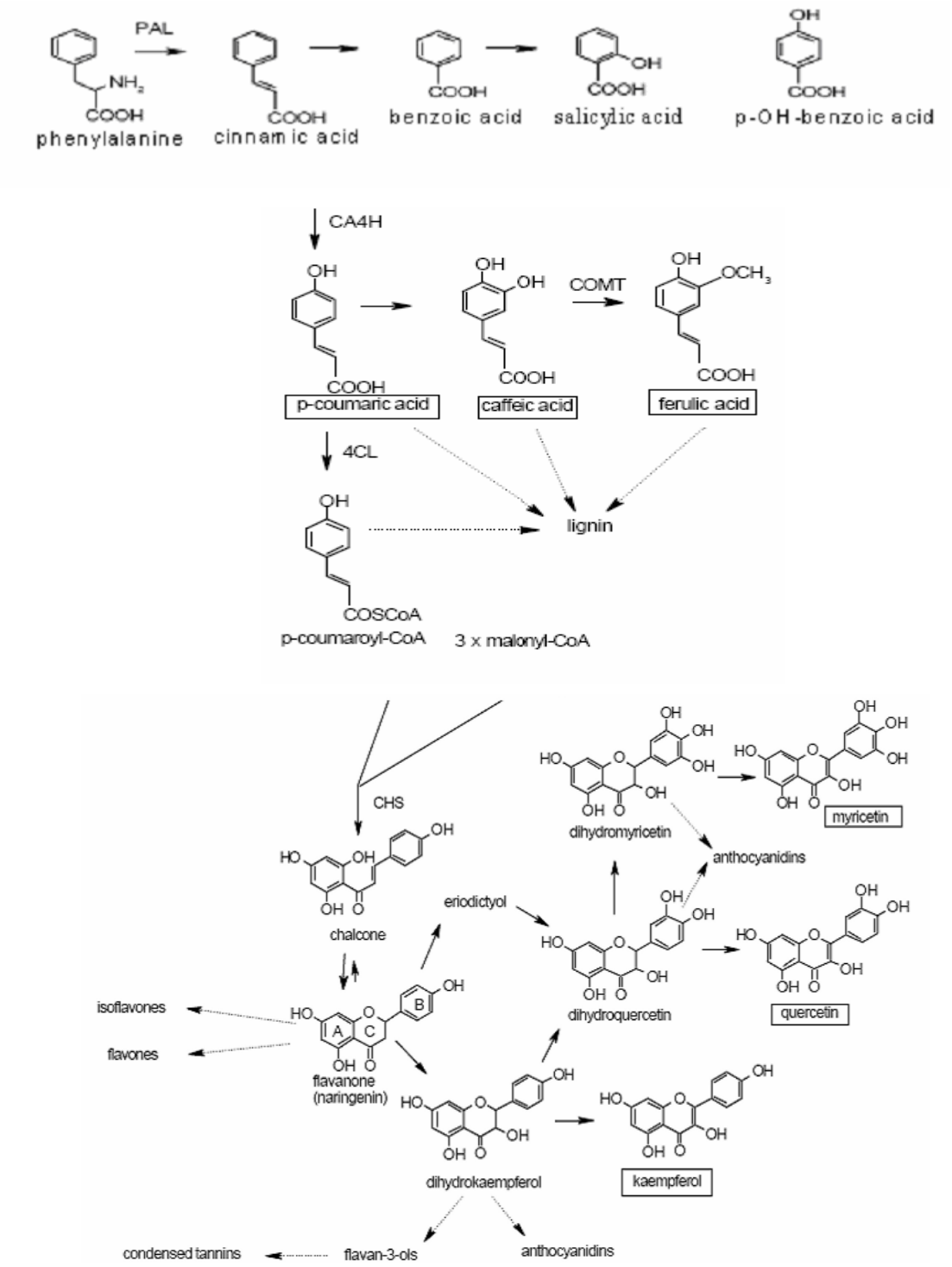


Fig. 1-1: Biosynthesis of hydroxycinnamic acids, hydroxybenzoic acids and flavonoids (Haddock et al. 1982, Harborne 1988, Hahlbrock and Scheel 1989, Maas et al. 1992, Bennet and Wallsgrove 1994, Dixon and Paiva 1995, Strack 1997). Solid arrows represent well-characterised reactions catalysed by single enzymes. Dashed lines represent transformations that require multiple enzymes that are less characterised, or vary among plant species. Enzymes: (CA4H), cinnamic acid 4-hydroxylase; (CHS), chalcone synthase; (4CL), 4-coumarate:coenzyme A ligase; (PAL), phenylalanine ammonia-lyase ; (COMT), caffeic acid O-methyltransferase.

The levels of the enzymes of phenylpropanoid metabolism are tightly and coordinately regulated during responses of plants to changes in their environment (Hahlbrock et al., 1976; Bolwell et al., 1985; Ni et al., 1996). In numerous plant species, accumulation of phenylpropanoid compounds was detected in response to different stresses. Cinnamate 4-hydroxylase is induced by light, elicitors, and wounding (Fahrendorf and Dixon, 1993; Buell and Somerville, 1995; Batard et al., 1997; Bell-Lelong et al., 1997). Its induction often is closely coordinated with that of PAL (Mizutani et al., 1997). As mentioned above, the phenylpropanoid biosynthetic pathway leads to the synthesis of a vast array of biologically active secondary metabolites, such as UV absorbing flavonoids, anthocyanin pigments, stilbene and isoflavonoid phytoalexins, and structural molecules such as lignin (Hahlbrock and Scheel, 1989; Smith, 1996). One important role of these compounds is the protection of plant tissues from damage by oxidative stress, wounding, and pathogen infections. Stress-induced activation of the biosynthesis of these compounds has been studied in many plant species (Dixon and Pavia, 1995). In addition to their protective roles, phenylpropanoids have an important effect on plant qualities such as texture, flavor, color, and processing characteristics. It has been argued for a long time that the enzymes of the phenylpropanoid pathway are organized in one or more enzyme complexes. Phenylpropanoid metabolism is largly unique to plants, and although quite complex one of the best-studied pathways in plants. There is now evidence for the coordinated expression of enzymes, channeling intermediates into a controlled metabolic flux which triggers the subsequent pathways in a concerted manner (Winkel, 2004).

1.3 Benzoic Acids

Benzoic acids serve important functions in plants. The aromatic metabolites are precursors of a number of natural products including many bioactive compounds. Benzoic acid is the biosynthetic building block of benzoyl and benzyl groups that are structural elements in numerous secondary products (Fig.1-2), such as zaragozic acids (Bergstrom et al., 1995; Baloglu et al., 1999), cocaine, paclitaxel, shikonin, cytokinins, and vanillin.

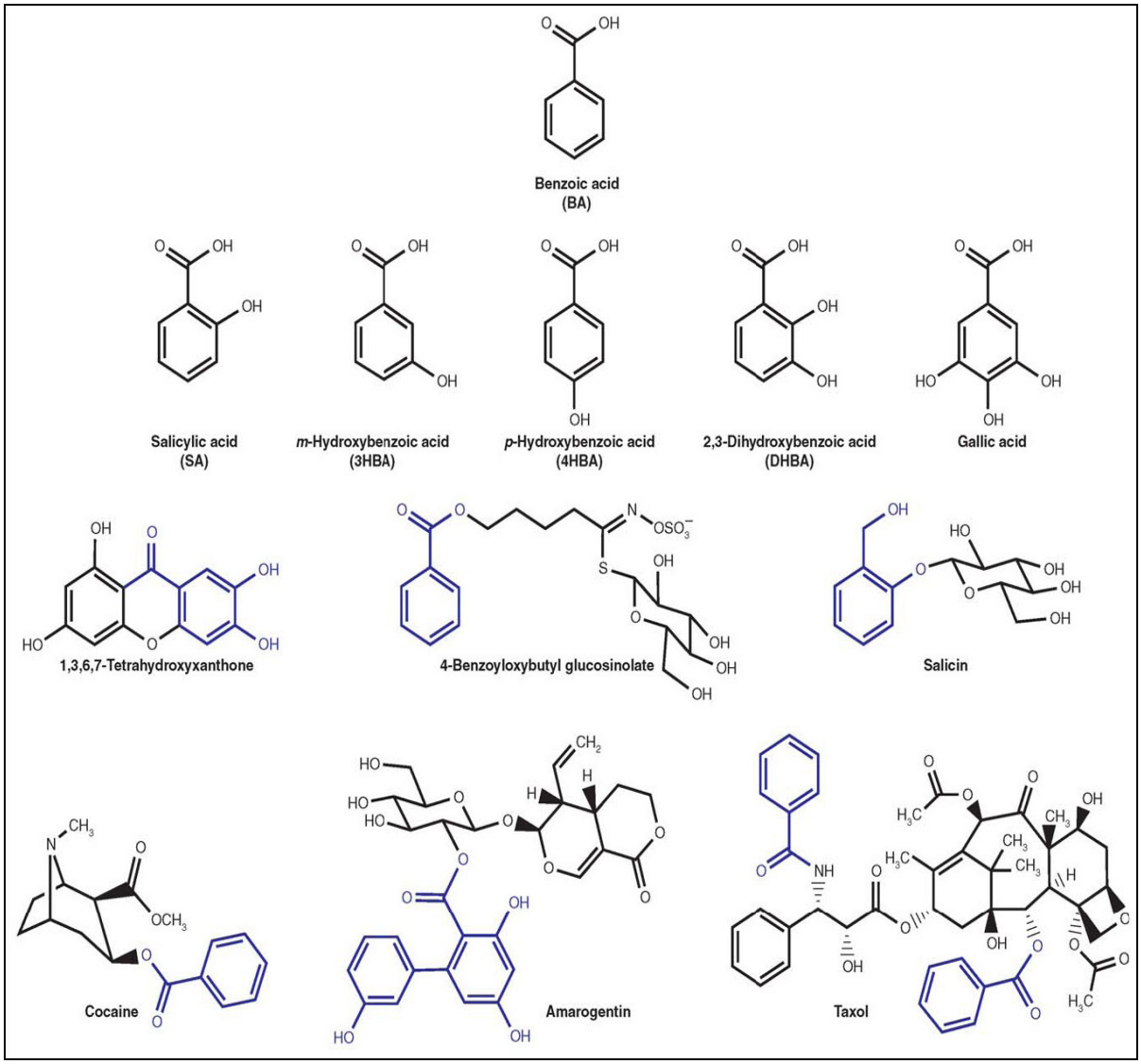


Fig. 1-2: Representative benzoic acids and secondary compounds that incorporate or are biosynthesized from a benzoyl moiety highlighted in blue

Cocaine is a member of the tropane alkaloids which possess the azabicyclo [3,2,1] octane system as a common structural element. The formation of the tropane nucleus follows a different route to chemical synthesis. It involves *N*-methylation and cyclization of ornithine-derived putrescine, followed by condensation with acetoacetate. Methylation of the free carboxylate group followed by ring closure and benzoylation of the stereospecifically reduced ketone group gives cocaine (Hopper, 1995). Benzoic acid which is required for the esterification step is synthesized from phenylalanine via cinnamic acid.

Paclitaxel (Taxol®) is a highly effective anticancer drug used widely in the treatment of various carcinomas, melanomas, and sarcomas (Goldspiel, 1997). It was originally isolated from *Taxus brevifolia* (Pacific yew) (Wani et al., 1971). Zamir et al. (1992) and Fett-Neto et al. (1994) demonstrated the increased accumulation of paclitaxel in the presence of the precursors phenylalanine, benzoic acid, serine, *N*-benzoylglycine and glycine in callus and cell cultures of *T. cuspidata*. Srinivasan et al. (1996) observed the same kind of effect with the addition of phenylalanine and benzoylglycine in cell cultures of *T. chinensis*. Fleming et al. (1993) reported the side chain of taxol to be formed via *β*-phenylalanine. The non-side chain taxane, baccatin III, has been proposed as the immediate precursor to paclitaxel via side chain addition at C-13 (Walker and Croteau, 2001). Taxane side chains are typically in the form of phenylisoserine, Winterstein's acid, or cinnamic acid. It was originally thought that side chain formation occurred by deamination of phenylalanine to form cinnamic acid through phenylalanine ammonia lyase (PAL) and subsequent reamination to form Winterstein's acid (Leete and Bodem, 1966).

Shikonin is formed from two precursors, 4-hydroxybenzoate (4HB) and geranyl diphosphate (GPP). GPP for shikonin biosynthesis is formed from mevalonate (Li et al., 1998), and 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR) has been identified as a regulatory enzyme in this part of the pathway (Lange et al., 1998). 3-Geranyl-4-hydroxybenzoate (GBA), the first specific intermediate in shikonin biosynthesis, is produced from GPP and 4HB under catalysis of 4HB geranyltransferase (Heide and Tabata, 1987). 4-Hydroxybenzoic acid also leads to the formation of ubiquinones. Furthermore, gallic acid (3,4,5-trihydroxybenzoic acid) serves as a fundamental precursor for gallotannins and ellagitannins (Dewick, 1997). Based mainly on radiolabeling studies, it was proposed that gallic acid could be formed from phenylalanine via caffeic acid (3,4,5-trihydroxycinnamic acid) or protocatechuic acid (Kato et al., 1986; Brucker, 1957). Recently, gallic acid was shown to be derived from 3-dehydroshikimic acid (Werner et al., 1997).

Natural cytokinins (CKs) are derivatives of adenine, with either an isoprenoid or a benzyl moiety substituted at the *N*-6 position, which is usually modified further. CKs exist in the form of free bases, ribosides, nucleotides, glyco-alanyl, and acetylconjugates (Jameson, 1994). Cytokinins substituted at *N*-6 with an aromatic benzyl or 3-hydroxybenzyl group are *N*-6-benzyladenine and *meta*-topolin, respectively. *meta*-Topolin has been used in different crops to induce axillary bud proliferation (Holub et al., 1998).

Vanillin, the most important flavor compound in the food and flavor industry and one of the oldest flavoring agents, is obtained from the bean-like fruit of the orchid *Vanilla planifolia* Andrews. It is believed that vanillin (4-hydroxy-3-methoxybenzaldehyde) is a product of the phenylpropanoid pathway (Havkin-Frenkel et al., 1999; Dignum et al., 2001). Vanillin is produced from 4-coumaric acid via 4-hydroxybenzaldehyde and 3,4-dihydroxybenzaldehyde (Havkin-Frenkel et al., 1996; Havkin-Frenkel et al., 1999). The first step from 4-coumaric acid to 4-hydroxybenzaldehyde is mediated by a putative 4-hydroxybenzaldehyde synthase (Podstolski et al., 2002).

1.3.1 Salicylic Acid and Systemic Acquired Resistance to Pathogens

Systemic **A**cquired **R**esistance (SAR) refers to an induced resistance that develops systemically following a prior challenge, typically with a necrotic-lesion forming pathogen, nonpathogen, chemical agent, or, in a few cases, an abiotic stressor (Hammerschmidt, 1999). SAR is dependent upon the action of salicylic acid (SA, 2-hydroxybenzoic acid), which is produced locally and systemically in the plant (Meuwly et al., 1995; Pieterse et al., 2002) and is effective against a fairly broad range of viral, eubacterial, and fungal pathogens. The fact that benzoic acid plays a central role in salicylic acid biosynthesis makes the understanding of this pathway even more important. SA is considered one of the key endogenous signals involved in the activation of numerous plant defense responses (Shah and Klessig, 1999). Early evidence showed that application of SA induced resistance against several pathogens and the expression of pathogenesis-related proteins in a variety of plants (White, 1979; Ward et al., 1991). These properties subsequently were found to mimic the natural defense response in tobacco (*Nicotiana tabacum*) and cucumber (*Cucumis sativa*) when elevated endogenous SA levels were correlated with induced resistance to the invading pathogen (Malamy et al., 1990; Métraux et al., 1990). Given the importance of SA in disease resistance, the pathway of SA biosynthesis may represent a major control point in plant defense responses. The biosynthetic pathway of SA appears to begin with the conversion of phenylalanine to *trans*-

cinnamic acid (*t*-CA) catalyzed by phenylalanine ammonia-lyase (PAL). *β*-Oxidation is the most likely mechanism underlying the early step of SA biosynthesis in tobacco (Ribnicky et al.,1998). The conversion of *trans*-cinnamic acid into SA has been proposed to proceed via chain shortening to produce benzoic acid, followed by hydroxylation at the C_2 position to derive SA (Yalpani et al., 1993). The latter step is likely to be catalyzed by a cytochrome P450 monooxygenase, called benzoic acid 2-hydroxylase, the activity of which is induced by either pathogen infection or exogenous benzoic acid application (León et al., 1993). Interestingly, benzoic acid and SA each can be conjugated to glucose, and regulation of SA levels through SA or benzoic acid conjugation may be important. In healthy tobacco plants, there is a large pool of conjugated benzoic acid that decreases transiently in size after pathogen infection. The decrease in conjugated benzoic acid levels correlates with an increase in free benzoic acid and SA (Yalpani et al., 1993). Once SA accumulates, it is rapidly converted to *β*-O-D-glucosylsalicylic acid (SAG), which apparently is not active in disease resistance (Lebn et al., 1993). Conversion of SAG to free SA represents another potential mechanism for increasing levels of free SA. The enzyme responsible for SA synthesis from benzoic acid has been shown to be induced by H_2O_2 (León et al., 1995). Draper (1997) proposed a model where initial pathogen infection would lead to rapid and short-living production of H_2O_2. Methyl salicylate (MeSA) may represent an airborne defense signal that activates resistance in nearby plants (Shulaev et al., 1997). MeSA is synthesized from SA and apparently acts by being converted back to SA (Seskar et al., 1998). MeSA is a constituent of the floral scent of *Clarkia breweri* flowers (Dudareva et al., 1998).

1.3.2 Benzoic Acids as Precursors of Benzophenone Biosynthesis

Benzophenones are the immediate precursors of xanthones in higher plants. Both classes of phenolic natural products include a number of pharmacologically active compounds. Their structures are related to those of chalcones and flavonoids, respectively. The biosynthesis of benzophenones and xanthones involves benzoic acids. Either benzoyl-CoA or 3-hydroxybenzoyl-CoA is subjected to sequential chain elongation by three acetate units from malonyl-CoA to yield an intermediate benzophenone (Fig.1-3) (Beerhues, 1996; Schmidt and Beerhues, 1997). The respective coenzyme A ligases catalyze the esterification of benzoic acid or 3-hydroxybenzoic acid by coenzyme A and have been detected in cell cultures of the medicinal plants *Hypericum androsaemum* (Clusiaceae) and *Centaurium erythraea* RAFN (Gentianaceae) (Barillas and Beerhues, 1997; Abd El-Mawla and Beerhues, 2002).

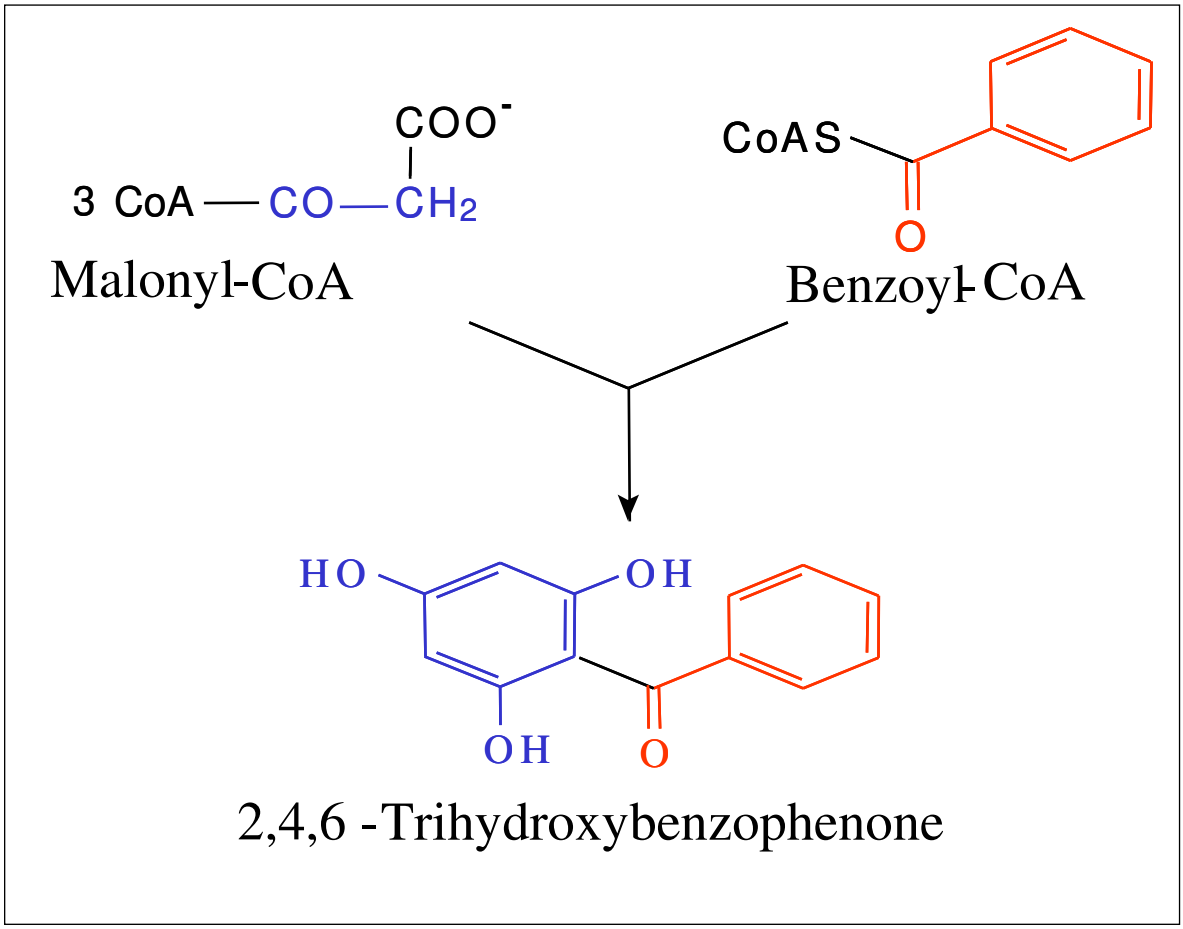


Fig. 1-3: Reaction catalyzed by benzophenone synthase (Schmidt and Beerhues, 1997)

Phlorbenzophenone is a central intermediate in benzophenone and xanthone metabolism. It is either polyprenylated to give complex molecules or intramolecularly cyclized to yield xanthones (Fig.1-4). In the first case, simple benzophenones can undergo stepwise prenylation using prenyl donors such as dimethylallyl diphosphate (Hu, et al., 2000). Accompanying cyclizations of the prenyl substituents lead to the formation of bridged polycyclic compounds that are widely distributed in Clusiaceae. The combination of challenging chemical structure and intriguing pharmacological activity makes the complex constituents attractive molecules for biotechnological research. For example, garcinol is a potent inhibitor of histone acetyltransferases both *in vitro* and *in vivo* and induces apoptosis (Balasubramanyam, et al., 2004). Sampsoniones that are characterized by unique caged tetracyclic skeletons include cytotoxic compounds whose mechanisms of action and target structures are under study (Hu, et al., 1999; Matsumoto, et al., 2003). Further polyprenylated polycyclic benzophenone

derivatives exhibit antibacterial and HIV-inhibitory activities (Cuesta, et al., 1999, Fuller, et al., 1999). Besides polyprenylation, benzophenones can undergo intramolecular cyclization to give xanthones (Fig.1-4). 2,3,4,6-Tetrahydroxybenzophenone is converted to 1,3,5- and 1,3,7-trihydroxyxanthones by regioselective oxidative phenol couplings that are catalyzed by cytochrome P450 enzymes (Peters et al., 1998).

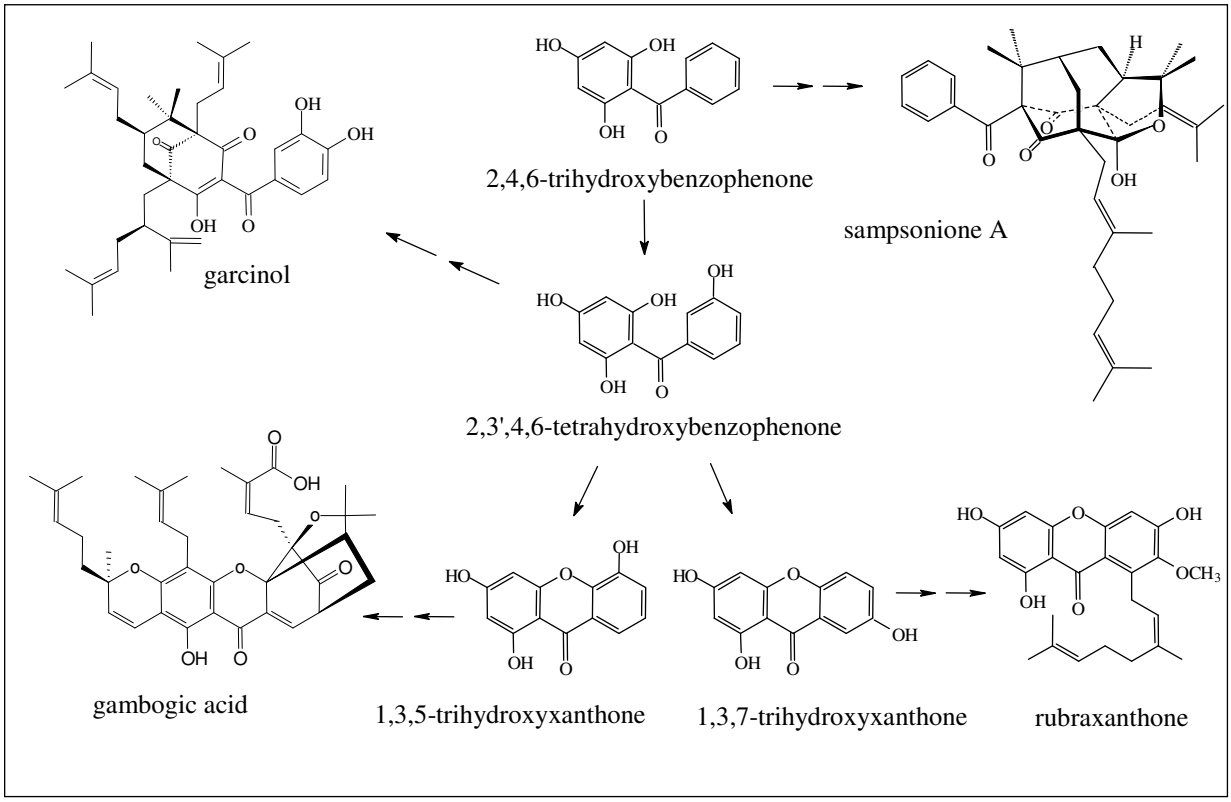


Fig. 1-4: Survey of benzophenone and xanthone metabolism

In cell cultures of *Hypericum androsaemum*, benzophenone synthase stepwise condenses one molecule of benzoyl-CoA with three molecules of malonyl-CoA to give a tetraketide intermediate that is cyclized by intramolecular Claisen condensation into 2,4,6-trihydroxybenzophenone (Fig.1-3) (Schmidt and Beerhues, 1997). BPS and CHS use the same mechanism of reaction, i.e. multiple decarboxylative condensations (Fig.1-4), but they use different starter substrates. Benzoyl-CoA and 4-coumaroyl-CoA arise biosynthetically from cinnamic acid, itself supplied by the PAL-catalyzed oxidative deamination of phenylalanine (Fig.1-5) (Liu et al., 2003). The subsequent 3'-hydroxylation is catalyzed by a cytochrome P450 monooxygenase. CHS supplies 4,2',4',6'-tetrahydroxychalcone for downstream enzymes that synthesize a diverse set of flavonoid compounds. Synthesis of chalcone by CHS involves the sequential condensation of one 4-coumaroyl-CoA molecule and three malonyl-CoA

molecules. After initial capture of the 4-coumaroyl moiety, each subsequent condensation step begins with decarboxylation of malonyl-CoA at the CHS active site.

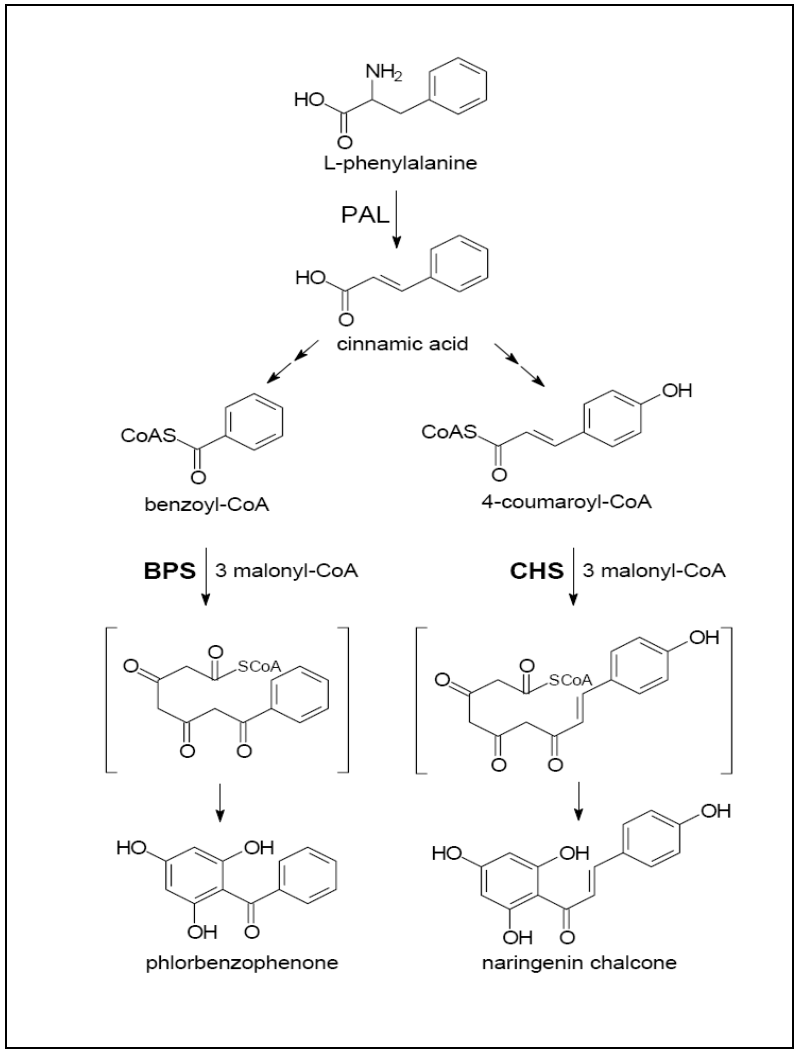


Fig. 1-5: Branching of the general phenylpropanoid pathway and the benzoic acid biosynthetic route leading to the formation of chalcone and benzophenone derivatives, respectively (Liu et al., 2003)

A number of plant polyketide synthases (PKSs) related to CHS by sequence identity, such as stilbene synthase (STS) (Schröder et al., 1998), bibenzyl synthase (BBS) (Preisig-Müller et al., 1995) and acridone synthase (ACS) (Junghanns et al., 1995), share a common reaction mechanism but differ from CHS in their substrate specificity and/or the stereochemistry of the polyketide cyclization reaction. For example, STS condenses one 4-coumaroyl- and three malonyl-CoA molecules, like CHS, but synthesizes resveratrol through a structurally distinct cyclization intermediate. The formation of 2,3',4,6-tetrahydroxybenzophenone was shown in cell-free extracts from cultured cells of *Centourium erythraea.* The enzyme is a benzophenone

synthase catalyzing the sequential condensation of three molecules of malonyl-CoA with one molecule of *m*-hydroxybenzoyl-CoA. The preferred starter substrate for this benzophenone synthase is 3-hydroxybenzoyl-CoA, yielding 2,3',4,6-tetrahydroxybenzophenone, the substrate for the regioselective cyclizations. Therefore all xanthones and their derivatives in *C. erythraea* are 5-oxygenated, in contrast to the constituents in *H. androsaemum* which all carry a 7-hydroxy group. The last step in the biosynthesis of the tetrahydroxyxanthones is the introduction of a hydroxyl group in position 6. This reaction is carried out by a cytochrome P450-dependend monooxygenase, xanthone-6-hydroxylase. Cell cultures of *C. erythraea* finally accumulate 3,5,6,7,8-pentamethoxy-1-*O*-primeverosylxanthone (Beerhues, 1994; Beerhues, 1996; Peters et al., 1998; Schmidt et al., 2000). Cell cultures of *H. androsaemum* contain a number of prenylated xanthone derivatives (Schmidt et al., 2000) and were therefore selected as starting material for studies reported here.

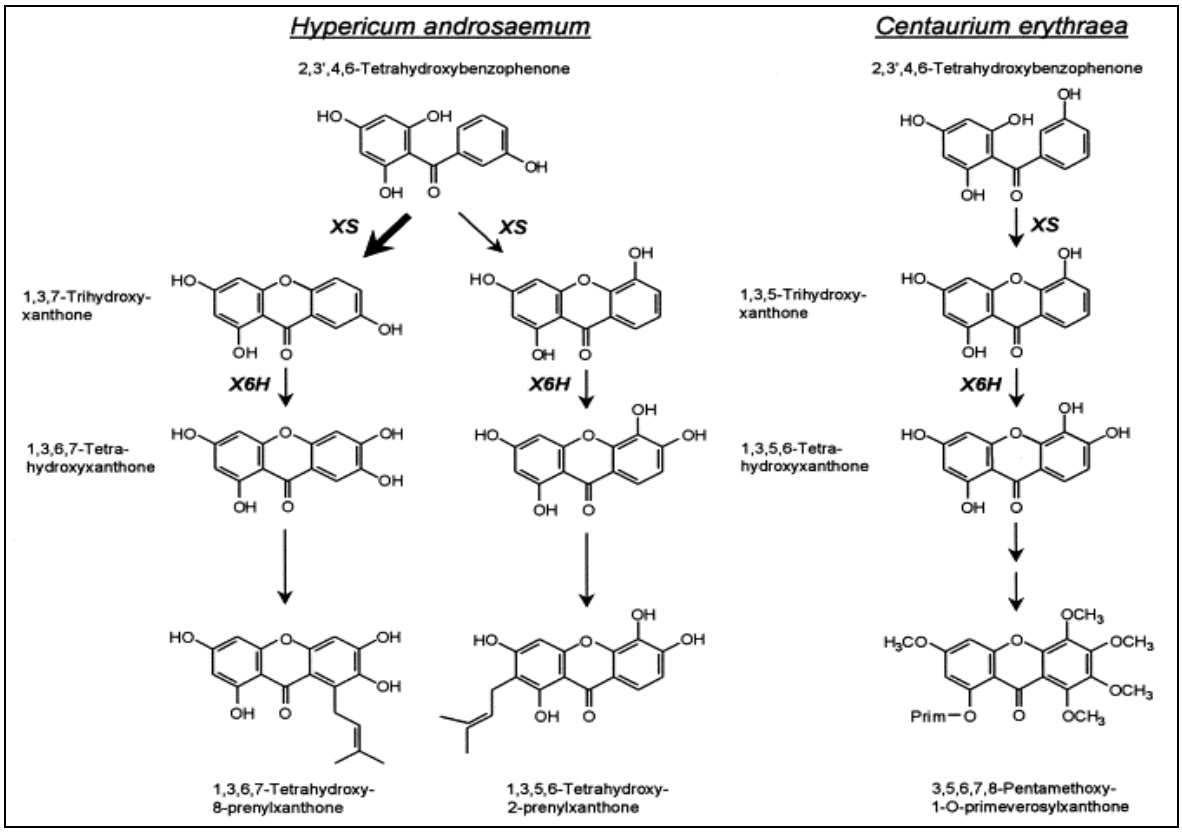


Fig. 1-6: Xanthone biosynthesis in cell cultures of *H. androsaemum* and *C. erythraea*

1.4 *Hypericum androsaemum* (Tutsan)

Hypericum androsaemum L. (Clusiaceae), commonly known as "Tutsan", got the latter name from the French "toute saine", meaning to heal all. There is certainly substantial folk history of the leaves being used to heal wounds. Tutsan is a medicinal plant species growing in West Europe (Flora Europaea, 1968). Its leaves have been employed in the preparation of an infusion used in folk medicine for its diuretic and hepatoprotective activities (Costa, 1994). When freshly picked, the tongue-shaped leaves have little smell, but when dried for a few days they develop a strong aromatic odour. The few recent reports on the chemical characterization of this species mainly concern the composition of the phenolic fraction. Phenolic acids and flavonoids (Seabra et al., 1992; Dias et al., 1999) as well as xanthones (Nielsen et al., 1979) and xanthone *C*-glucosides (Kitanov et al., 1998) have been identified in *H. androsaemum* plants. The essential oil of the species is less studied. The *n*-alkanes nonane and undecane as well as the monoterpenes α-pinene, β-pinene, myrcene, limonene, geraniol, and α-terpineol were identified as essential oil components (Mathis et al., 1964; Nogueira et al., 1998). The presence of α-terpineol and hydrocarbon waxes ($C_{19}H_{40}$, $C_{21}H_{44}$, and $C_{23}H_{48}$) in *H. androsaemum* unripened seed capsules was also reported (Carnduff et al., 1966). Less is known on metabolites produced by in vitro cultures of this species. Phenolic compounds, such as flavonoids, phenolic acids and xanthones, were identified in callus and cell suspension cultures (Dias et al., 1999; Dias et al., 2000; Schmidt et al., 2000; Valentão et al., 2003). The traditional therapeutic effects of this species can be attributed, at least partially, to the presence of these kinds of constituents. Phenolic compounds which are widely distributed in plants have been considered to play an important role as dietary antioxidants for the prevention of oxidative damage in living systems (Halliwell et al., 1995; Brown and Rice Evans, 1998; Harborne and Williams, 2000; Shi et al., 2001). *H. androsaemum* infusion exhibited antioxidant activity in vitro which was in part attributed to its phenolic composition, namely to caffeoylquinic acids and quercetin derivatives (Valentão et al., 2002).

Fig. 1-7: *Hypericum androsaemum* L. (Clusiaceae)

The native range of the species is Western Europe, scattered in southern Europe and Caucasia, Asia (northern Iran, western Syria), and northwest Africa (Davis, 1967). The leaves are oval-shaped and have an aromatic smell when crushed. The leaves are covered with pellucid dots. Flowers are bright yellow and arranged in terminal clusters. The plant is in flower from June to August. The scented flowers are hermaphrodite (have both male and female organs), and are pollinated by insects. The fruits are black berries, and the seeds ripen from August to September. *H. androsaemum* is a prolific seeder and can be propagated from seeds and cuttings.

1.5 Benzoic Acids as Precursors of Biphenyl Biosynthesis

Biphenyls and dibenzofurans are the phytoalexins of the Maloideae (Kokubun and Harborne, 1995; Hrazdina, 2003). The key enzyme of their biosynthesis is biphenyl synthase (BIS) which is a novel type III PKS (Liu et al., 2006). The homodimeric enzyme catalyzes the formation of the C_{12} skeleton of the two classes of compounds. In *Sorbus aucuparia* cell cultures, BIS transcripts were rapidly, strongly and transiently induced by elicitor treatment, resulting in a massive accumulation of aucuparin, the best known biphenyl phytoalexin (Liu et al., 2004). Activation by elicitation of genes encoding key enzymes of phytoalexin biosynthetic pathways such as isoflavonoid and furanocoumarin biosyntheses has previously

been reported, with the induction kinetics resembling that of BIS (Habereder et al., 1989; Parker et al., 1991). The preferred starter substrate for BIS is benzoyl-CoA that is a rare starter molecule for PKSs. In bacteria, soraphen A, enterocin and the wailupemycins originate from benzoyl-CoA as reaction primer (Moore et al., 2002). The first plant PKS found to use benzoyl-CoA as starter substrate was benzophenone synthase (BPS) (Liu et al., 2003). Both BIS and BPS catalyze the iterative condensation of benzoyl-CoA with three acetyl units from malonyl-CoA to give identical linear tetraketides (Fig.1-8). While BPS cyclizes this intermediate via an intramolecular C6→C1 Claisen condensation, BIS catalyzes an intramolecular C2→C7 aldol condensation and decarboxylative elimination of the terminal carboxyl group. The product of the BIS reaction is 3,5-dihydroxybiphenyl. The two type III PKSs share 54% amino acid sequence identity over their approx. 400 amino acids. Usually, plant type III PKSs share 60-95% identity with each other (Austin and Noel, 2003).

Cell suspension cultures of *S. aucuparia* were established that respond to yeast extract treatment with the accumulation of the biphenyl phytoalexin aucuparin (Liu et al., 2004). Recently, cDNA cloning, functional expression and characterization of BIS was reported (Liu et al., 2006). For comparison, chalcone synthase (CHS) that forms the carbon skeleton of the ubiquitous flavonoids from 4-coumaroyl-CoA as starter molecule was also cloned from *S. aucuparia* cell cultures.

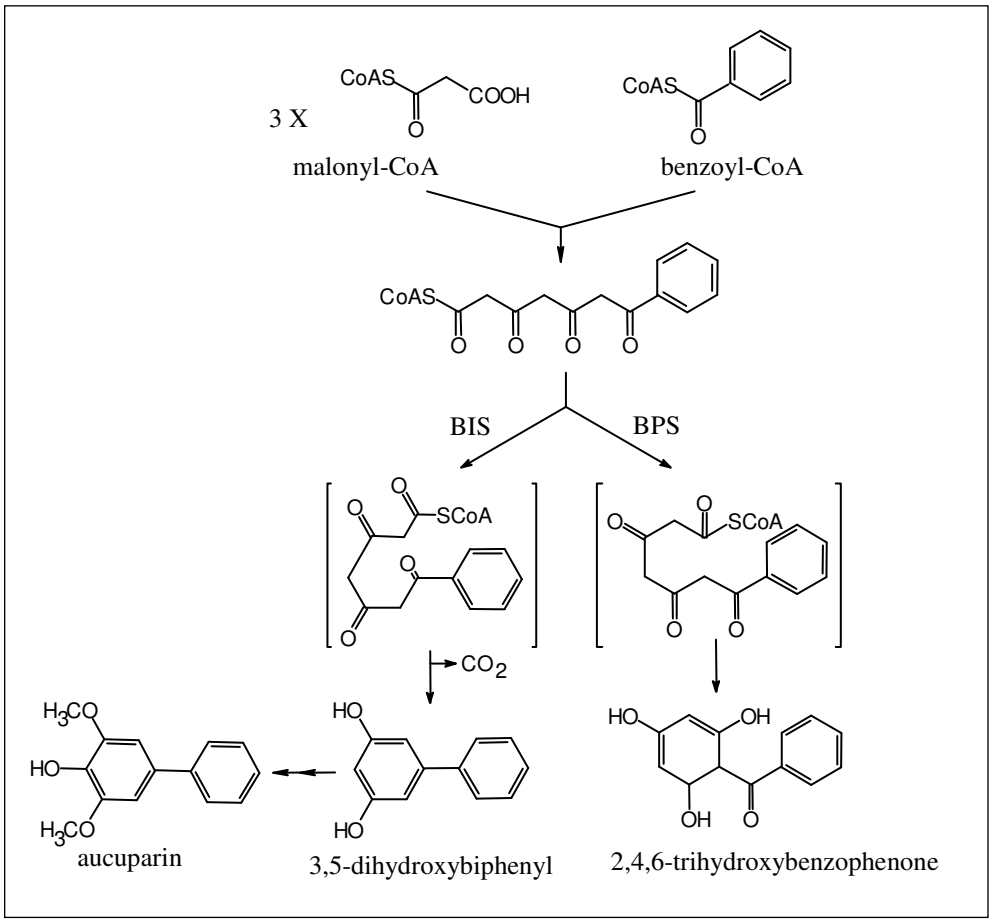


Fig. 1-8: Reactions of biphenyl synthase (*BIS*) and benzophenone synthase (*BPS*). The phytoalexin aucuparin results from subsequent reactions (Liu et al., 2004)

1.6 *Sorbus aucuparia* (mountain ash)

Sorbus aucuparia L., commonly known as European mountain ash, also called rowan, belongs to the family Rosaceae. Its growth habit is either tree or shrub. It is a perennial plant, which means it lives for more than two years, retaining its leaves in winter or not. The small white flowers are born in much-branched inflorescences and the fruits are red. Mountain ash belongs to the most unpretentious species as it grows on heavily degraded soil (Aas and Holdenrieder, 1992; Weihs, 1995). The bark is astringent and used in the treatment of diarrhoea and as a vaginal injection for leucorrhoea (Grieve et al., 1984). The fruit is antiscorbutic and astringent (Grieve et al., 1984; Chiej et al., 1984). The seeds contain cyanogenic glycosides (amygdalin) which, after release of the glycosyl moiety, produce the

extremely toxic prussic acid (Chevallier et al., 1996). In small quantities this acts as a stimulant to the respiratory system but in larger doses it can cause respiratory failure and death. It is therefore best to remove the seeds when using the fruit medicinally or as a food (Chevallier et al., 1996). Fatty oil is obtained from the seed (up to 22%) (Triska et al., 1975). A cosmetic face-mask is made from the fruits and is used to combat wrinkled skin (Chiej et al., 1984). A black dye is obtained from the young branches (Komarov, 1968). Mountain ash fruits contain cryptoxanthin, carotene, flavonoids (quercetin, iso-quercetin and rutin), E and B vitamins, anthocyans, tannic substances, phospholipids, pectin substances, sorbitol, catechins, saponins, various sugars as well as malic, citric, dihydroxysuccinic and parasorbic acids.

Fig. 1-9: *Sorbus aucuparia* L. (Rosaceae)

Previous studies showed that the biphenyl aucuparin and its derivatives which occur in the sapwood tissue of *S. aucuparia* are essentially absent from healthy tissue and only produced as phytoalexins following fungal infection (Kokubun et al., 1995). The accumulation of aucuparin is a typical induced defense reaction in response to microbial attack (Bailey and Mansfield, 1982; Brooks and Watson, 1985, Pedras and Khan, 2000; Pedras et al., 2000). Aucuparin was also identified as phytoalexin in the leaves of *S. aucuparia* (Kokubun and Harborne, 1994). Furthermore, it was found together with its 2` and 4`-oxygenated derivatives in the sapwood of *Aronia, Chaenomeles, Eriobotrya, and Malus* species following fungal inoculation or natural infection (Kokubun and Harborne, 1995). Aucuparin and the 1,3,7-

trihydroxy-2-(3-methylbut-2-enyl)-xanthone showed antimicrobial activity against *B. subtilis* and similar results were obtained with strains from clinical specimens. Penicillin-resistant *Staphylococcus aureus* (PRSA) strains seemed to be more sensitive to aucuparin and 1,3,7-trihydroxy-2-(3-methylbut-2-enyl)-xanthone than penicillin-susceptible *S. aureus* (PSSA) (Diógenes, 2002).

1.7 Biosynthesis of Benzoic Acids in *Hypericum androsaemum* Cell Cultures

Despite its simple structure and widespread occurrence, the biosynthesis of benzoic acid and its CoA-thioester is only partially understood. In the field of plant secondary metabolism, two major routes from L-phenylalanine to benzoic acid have been reported (Fig. 1-10): the *β*-oxidation-type pathway including four CoA-ester intermediates and, alternatively, the non-*β*-oxidative pathway via benzaldehyde as a key intermediate (French et al., 1976; Yazaki et al., 1991; Schnitzler et al., 1992). The contradictory results concerning the metabolic pathways leading to benzoic acid suggest that several side-chain shortening routes might exist in benzoic acid biosynthesis in different plants and even in the same plant depending on the physiological conditions. Both routes possess common intermediates before branching to involve either oxidation, thiolation or retro-aldol cleavage followed by oxidation.

In cell cultures of *H. androsaemum,* benzoic acid is formed from cinnamic acid, as shown by feeding experiments with radiolabeled precursors (Abd El-Mawla et al., 2001). The feeding experiments were performed with [U-^{14}C] L-phenylalanine, [7-^{14}C] benzoic acid and [7-^{14}C] 3-hydroxybenzoic acid. All three precursors were efficiently incorporated into an elicited xanthone in *H. androsaemum*, whereas 3-hydroxybenzoic acid was the only precursor to be incorporated into xanthones in *C. erythraea* cell cultures. In addition, an appreciable increase in phenylalanine ammonia-lyase activity occurred only in methyl-jasmonate-treated cell cultures of *H. androsaemum* (Beerhues, 1996; Schmidt and Beerhues, 1997; Abd El-Mawla et al., 2001).

The mechanisms underlying shortening of the side chain by a C2 unit are as follows. The CoA-dependent and *β*-oxidative mechanism is analogous to that underlying *β*-oxidation of fatty acids. It involves activation of cinnamic acid, hydration of the cinnamoyl moiety, oxidation of the hydroxyl group to a ketone and cleavage of the *β*-keto thioester via a reverse Claisen reaction. The CoA-independent and non-*β*-oxidative pathway involves hydration of

the free acid, side chain degredation via a reverse aldol reaction and oxidation of the intermediate aldehyde. In cell cultures of *H. androsaemum*, a combination of these two routes has been detected, as shown in (Fig.1-10) as CoA-dependent and non- β -oxidative pathway.

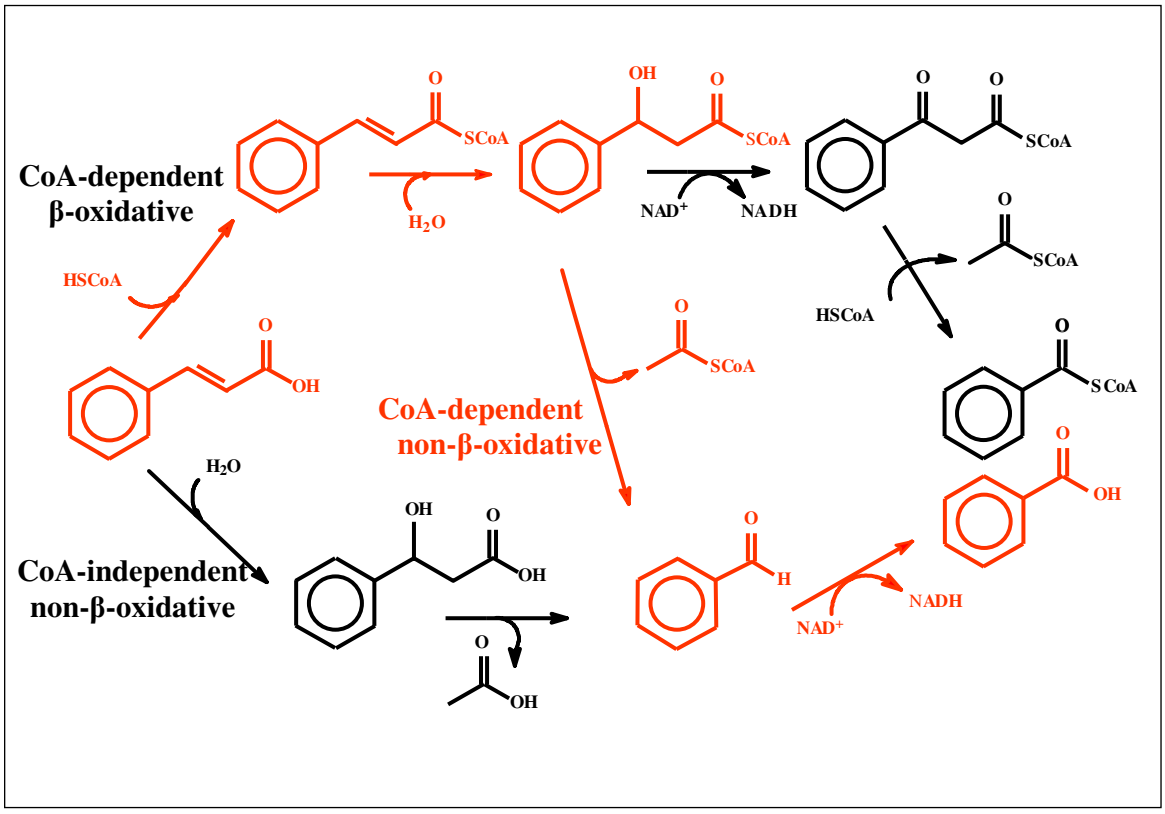


Fig. 1-10: Benzoic acid biosynthetic pathways via cinnamic acid. The pathway detected in *H. androsaemum* cell cultures is highlighted in red (Abd El-Mawla et al., 2001).

The complete sequence of enzymes participating in this route has been detected (Abd El-Mawla et al., 2001). The pathway involves cinnamoyl-CoA hydratase/lyase and benzaldehyde dehydrogenase. In addition, three CoA-ligases were separated by anion exchange chromatography. The three enzymes are cinnamate:CoA ligase, benzoate:CoA ligase, and 4-coumarate:CoA ligase. They differ clearly in their substrate specificities. Cinnamate:CoA ligase catalyzes the first step of benzoic acid biosynthesis and channels cinnamic acid from the general phenylpropanoid pathway into the benzoic acid biosynthetic pathway. Benzoate:CoA ligase is the first enzyme of xanthone and biphenyl biosyntheses and provides benzophenone synthase and biphenyl synthase, respectively, with the starter substrate for chain extension by C2 units from malonyl-CoA. 4-Coumarate:CoA ligase is, as mentioned above, a well-known enzyme of the general phenylpropanoid pathway and supplies the

substrate for the biosynthesis of a large number of natural products such as flavonoids and lignin.

The second step of benzoic acid biosynthesis in *H. androsaemum* cell cultures is catalyzed by cinnamoyl-CoA hydratase/lyase. This bifunctional enzyme catalyzes the addition of water to the double bond of the cinnamoyl residue and the subsequent cleavage of the resulting 3-hydroxy-3-phenylpropionyl-CoA. This putative intermediate was not detected in the enzyme assays, which argues against the possibility that two separate enzymes catalyze the consecutive reactions, hydration and cleavage. A gene encoding feruloyl-CoA hydratase/lyase has been isolated from a strain of *Pseudomonas fluorescens* (Gasson et al., 1998). The gene product catalyzes the metabolism of feruloyl-CoA to vanillin and acetyl-CoA. The hydrated derivative of feruloyl-CoA was also readily converted by the overexpressed enzyme. Thus, the CoA-dependent and non-*β*-oxidative pathway also underlies bacterial vanillin formation. In the plant species *Vanilla planifolia*, vanillic acid biosynthesis was found to proceed via isoferulic acid and 3,4-dimethoxycinnamic acid, which requires a final demethylation step (Funk and Brodelius, 1992).

The last reaction of benzoic acid biosynthesis is catalyzed by benzaldehyde dehydrogenase. Related enzyme activities have previously been detected in cell cultures of *Lithospermum erythrorhizon* and *Daucus carota*, both of which produce 4-hydroxybenzoic acid (Yazaki et al., 1991; Schnitzler et al., 1992). Accordingly, the dehydrogenases from these two species oxidized most efficiently 4-hydroxybenzaldehyde. In contrast, the preferred substrate of the *H. androsaemum* enzyme is benzaldehyde, which is in accordance with the substrate specificities of the two preceding biosynthetic enzymes. The second-best substrate of both the benzoic acid and the xanthone biosynthetic enzymes is the respective 3-hydroxylated derivative. This agrees well with previously proposed alternative routes of xanthone biosynthesis that start with either benzoic acid or 3-hydroxybenzoic acid (Schmidt and Beerhues, 1997). The activities of the benzoic acid biosynthetic enzymes were stimulated co-ordinately by methyl jasmonate. Furthermore, benzophenone synthase exhibited similar changes in its activity during cell culture growth and after methyl jasmonate treatment (Abd El-Mawla et al., 2001*;* Schmidt and Beerhues, 1997). Methyl jasmonate induces the accumulation of 1,3,6,7-tetrahydroxy-8-prenylxanthone and causes an increase in the level of PAL activity in *H. androsaemum* cell cultures (Abd El-Mawla et al., 2001).

1.8 Research Strategies and Objective

The objective of this work was to study, at the molecular genetic level, benzoic acid biosynthesis in elicitor-treated cell cultures of *Hypericum androsaemum* and *Sorbus aucuparia*. The focus was on the three CoA ligases previously detected at the biochemical level. The strategy involved isolation of mRNA from both cell cultures, RT-PCR, 5`- and 3`-RACE techniques and alignment of full-length sequences with databank entries. Studies were completed by functional expression in *E.coli* and characterization of the recombinant proteins.

2. MATERIALS

2.1 Plant Material

2.1.1 *Hypericum androsaemum*

Tissue cultures of *H. androsaemum* L. were grown as described previously (Peters et al., 1989; Beerhues and Berger, 1994). Callus tissue grown on solid LS medium was transferred to liquid LS medium and cultivated in the dark. The resulting cell suspension cultures were shaken in 300-ml Erlenmeyer flasks at 100 rpm and 25°C. Four g cells were inoculated into fresh LS medium (50 ml) at 7-day-intervals. Cultured cells were collected by suction filtration through Miracloth filter paper in a Büchner funnel.

2.1.2 *Sorbus aucuparia*

Callus of *Sorbus aucuparia* L. was derived from young shoots. After surface sterilization, stem segments were put on solid LS medium (Linsmaier and Skoog, 1965). Friable callus was used to initiate cell suspension cultures which were grown in liquid LS medium at 25°C in the dark on an orbital shaker at 100 rpm. Cultured cells (3 g) were transferred into 50 ml fresh medium every 14 days. Five-day-old cell cultures from the linear growth phase were treated with yeast extract (3 g l^{-1}). 9 h after the onset of treatment, cultured cells were used for the preparation of cell-free extract and the isolation of RNA.

2.2 Chemicals

Commonly used chemicals and gases were purchased in standard grade from the following companies:

Aldrich	Fluka	Roth
Applichem	Merck	Serva
Bio-Rad	Riedel-deHaën	Sigma

Aqueous solutions were prepared with deionized water that was purified by using the Milli-Q Water Purification system (Millipore).

2.2.1 Special Chemicals

2.2.1.1 Chemicals for Enzyme Assays

Benzoyl-CoA	Sigma Aldrich
ATP (Adenosine triphosphate)	Sigma Aldrich
Coenzyme A	Sigma Aldrich
2-Hydroxycinnamic acid	Sigma-Aldrich
3-Hydroxycinnamic acid	Sigma-Aldrich
Benzoic acid	Sigma-Aldrich
2-Hydroxybenzoic acid	Sigma-Aldrich
4-Hydroxycinnamic acid	Roth
Ferulic acid	Roth
Sinapic acid	Roth
Caffeic acid	Roth
Cinnamic acid	Merck
3-Hydroxybenzoic acid	Fluka
4-Hydroxybenzoic acid	Fluka
DMSO	Fluka
2-Hydroxybenzoyl-CoA	Laboratory collection
3-Hydroxybenzoyl-CoA	Laboratory collection
4-Hydroxybenzoyl-CoA	Laboratory collection
Caffeoyl-CoA	Laboratory collection
4-Coumaroyl-CoA	Laboratory collection
Feruloyl-CoA	Laboratory collection
Sinapoyl-CoA	Laboratory collection
Cinnamoyl-CoA	Laboratory collection
2-Hydroxycinnamoyl-CoA	Laboratory collection
3-Hydroxycinnamoyl-CoA	Laboratory collection
DTT (1,4-Dithiothreitol)	AppliChem

2.2.1.2 Chemicals for Gel Electrophoresis and Molecular Biology

Protein molecular weight ladder	Fermentas
PD10-sepharose G-25	Amersham Pharmacia Biotech.
Molecular weight marker kit	Sigma
Acrylamide / bis-acrylamide	BioRad
Ammonium persulfate (APS)	Bio-Rad
Bromphenol blue	Aldrich
Coomassie brilliant blue G250	Merck
Sodium dodecyl sulphate (SDS)	Roth
TEMED (*N,N,N´,N´*-Tetramethylethylenediamine)	BioRad
Agarose, NEE ultra quality	Roth
Formaldehyde, 37%	Roth
Agar-agar	Roth
D-glucose	Fluka
Methyl jasmonat	Serva
Select yeast extract	Life Technologies
Select peptone	Life Technologies
IPTG	Sigma Aldrich
X-Gal	Sigma Aldrich
Chloramphenicol	Roth
Ampicilline	Roth

2.3 Equipments

Balances	Large and small scale	Sartorius
Water bath	Typ 3041	Köttermann
	Exatherm U3	Julabo
Heater block	Dri-Block® DB 30	Techne
Incubator	HT	Infors
Centrifuge	Universal 32R	Hettich
	Biofuge 13	Heraeus Sepatech
	Biofuge pico	Heraeus

HPLC	1525 Binary HPLC Pump	Waters
	2487 Dual λ Absorbance Detector	Waters
	Breeze GBC Software	Waters
FPLC	Biologic-System	Bio Rad
	Controller	BioRad
	Fraction collector 2128	BioRad
Photometer	Ultraspec 1000	Pharmacia Biotech
Power supply	Standard Power Pack P25	Biometra
	Power Pack 300	Bio Rad
Electrophoresis	Mini-sub® Cell GT	BioRad
	Wide Mini- sub® Cell GT	BioRad
Electrophoresis	Protein-Chamber	Biometra; BioRad
Vacuum concentrator	Genelac SF 50	Biometra
pH-Meter	Digital pH Meter 325	WTW (Wissenschaftlich-Technische-Werkstätten)
Autoclave	Vapoklav	Sterilco
Clean bench	LaminAir HLB 2472	Heraeus
Freeze dryer	Alpha	Christ
Water distillation Unit	Milli-Q Reagent Water	Millipore
Microwave oven	Optiquick	Moulinex
Magnetic rotator	Vibrax-VXR	IKA–Labortechnik (Janke & Kunkel), Staufen
	VF2	
	Combimag Ret	
Multiphor™ II Electrophoresis unit		Pharmacia

2.4 Nutrient Media for Plant Tissue Cultures

The prepared medium was distributed into vessels such as Erlenmeyer flasks (for example, 50 ml aliquots in 300 ml Erlenmeyer flasks) and sealed with aluminium caps sterilized by using an autoclave at 120°C for 20 minutes.

Tab. 2-1: Stock solutions and preparation of media

Medium	Stock Solution Ingredients		Preparation and storage.
LS-Liquid medium (Linsmaier and Skoog,1965)	**I**		Mix with 800 ml distilled water, add and dissolve sucrose, adjust the volume to one liter with water and the pH to 6-6.3 with concentrated NaOH solution. Autoclave.
	KNO_3	1900 mg/l	
	NH_4NO_3	1650 mg/l	
	$CaCl_2 \cdot 2H_2O$	440 mg/l	
	$MgSO_4 \cdot 7H_2O$	370 mg/l	
	KH_2PO_4	170 mg/l	
	II		
	$MnSO_4 \cdot H_2O$	16.90 mg/l	
	$ZnSO_4 \cdot 7H_2O$	10.60 mg/l	
	KI	0.83 mg/l	
	H_3BO_3	6.20 mg/l	
	$Na_2MoO_4 \cdot 2H_2O$	0.25 mg/l	
	$FeSO_4 \cdot 7H_2O$	27.80 mg/l	
	$CuSO_4 \cdot 5H_2O$	0.025 mg/l	
	$CoCl_2 \cdot 6H_2O$	0.025 mg/l	
	III		
	Titriplex III (Na_2 EDTA . $2H_2O$)	41.30 mg/l	
	IV		
	myo – Inositol	100.0 mg/l	
	Thiamine hydrochloride	0.4 mg/l	
	V		
	2,4-Dichlorphenoxy-acetic acid	0.22 mg/l	
	1-Naphtylacetic acid	0.186 mg/l	
	VI		
	sucrose	30 g/l	
LS-Solid medium	See LS-medium, but with adding 8.0 g/l select agar and heating the medium to nearly 80 °C		Nutrient medium for callus cultures

To solidify LB medium, 1.2% agar were added after adjusting the pH of the medium to 7.5 with NaOH. The medium was then sterilized by autoclaving. Filter-sterilized antibiotics were added to agar plates or into liquid medium just before use.

2.5 Solutions and Buffers for Biochemical Analysis

2.5.1 Buffers for Extraction and Enzyme Incubation

Tab. 2-2: Buffers for Extraction and Enzyme Incubation

Name	Ingredients	Preparation and storage
Extraction buffer	0.1 M Potassium dihydrogen phosphate 1.36 g Water ad 100 ml	pH adjusted by KOH, autoclaved, and stored at 4°C.

2.5.2 Solution for Regeneration of PD_{10}-Columns (Amersham Biosciences)

Tab. 2-3: Solution for Regeneration of PD_{10}-Columns

Name	Ingredients	Preparation and storage
NaOH cleaning agent	NaOH 0.16 M	Wash the column with five column volumes of 0.16 M NaOH and then five column volumes of distilled water

2.5.3 Buffer for Protein Determination

Tab. 2-4: Buffer for Protein Determination

Name	Ingredients	Preparation and Storage
Bradford-dye solution	Coomassie®-brilliant blueG-250 100 mg Ethanol 96 % 50 ml Orthophosphoric acid 100 ml Water ad 1000 ml	Dissolve Coomassie® -Brilliant blue G-250 in ethanol, add orthophosphoric acid and complete the volume with water. The solution should be stored in the refrigerator and filtered before use.

2.5.4 Buffers and Solutions for Gel Electrophoresis (SDS-PAGE)

Tab. 2-5: Buffers and Solutions for Gel Electrophoresis (SDS-PAGE)

Name	Ingredients	Preparation and Storage
Stacking gel	Distilled water 3.4 ml 1 M Tris-HCl (pH 6.8) 0.63 ml Acrylamid/bis 30 % 0.83 ml 10 % (w/v) SDS 0.05 ml 10 % (w/v) APS 0.05 ml TEMED 5 µl	
Separating gel (12 %)	Distilled Water 3.3 ml 1.5 M Tris-HCl (pH 8.8) 2.5 ml Acrylamid/Bis 30 % 4.0 ml 10 % (w/v) SDS 0.1 ml 10 % (w/v) APS 0.1 ml TEMED 4 µl	
2x Protein loading buffer	Distilled Water 2.7 ml 0.5 M Tris-HCl (pH 6.8) 1.0 ml Glycerin 2.0 ml 10 % (w/v) SDS 3.3 ml β-Mercaptoethanol 0.5 ml 0.5 % (w/v) Bromphenolblue 0.5 ml	Store at 4 °C
10x SDS-electrode buffer	Tris base 15 g Glycin 72 g Na-SDS 5 g Water up to 500 ml	Store at 4 °C, dilute 1:10 before use
Staining solution	Coomassie-blue R 250 25 ml Methanol 100 ml Acetic acid 20 ml Water up to 200 ml	
Coomassie-blue	Coomassie-blue R 250 0,5 g Water up to 50 ml	Dissolve the dye in water and filtrate
Destaining solution	Methanol 30 ml Acetic acid 20 ml Distilled water up to 200 ml	
Silver staining Fixing solution	30% ethanol 150 ml 10% acetic acid 50 ml Distilled water up to 500 ml	Store at room temperature

Sensitizing solution	30 % ethanol 150 ml 0.5 M sodium acetate 20.5 g 0.2 % $Na_2S_2O_3$ 1 g Distilled water up to 500 ml	Store at room temperature
Silver solution	0.2 % silver nitrate 0.6 g 0.01 % formaldehyde 30 µl Distilled water up to 300 ml	Prepare freshly
Developing solution	6 % Na_2CO_3 18 g 0.02 % formaldehyde 60 µl Distilled water up to 300 ml	Prepare freshly
Stop solution	1.5 % Na_2EDTA 4.5 g Distilled water up to 300 ml	

2.5.5 Buffers and Solutions for Protein Purification

2.5.5.1 Buffers for Affinity Chromatography

Tab. 2-6: Buffers for Affinity Chromatography

Name	Ingredients	Preparation and storage
Binding buffer pH 7.4	20mM Sodium phosphate 1.248 g 0.5M NaCl 11.688g 80 mM Imidazole 1.053g Water up to 400 ml	Filter, degas and store at 4°C, adjust pH by adding concentrated HCl.
Elution buffer pH 7.4	20mM Sodium phosphate 1.248 g 0.5 M NaCl 11.688g 1 M Imidazole 12.053g Water up to 400 ml	Filter, degas and store at 4°C, adjust pH by adding concentrated HCl.

2.5.5.2 Buffer for Gel Filtration (HiPrep16/60 SephacryS-200)

Tab. 2-7: Buffer for Gel Filtration

Name	Ingredients	Preparation and storage
Elution buffer pH 7.5	50 mM Tris 6.057g 100 mM KCl 7.456 g 1 mM DTT 1.5 mg Water up to 1000 ml	Filter, degas, and store at 4°C.

2.6 Buffers and Media for Molecular Biology

2.6.1 Media for the Cultivation of *E. coli*

Tab. 2-8: Media for the Cultivation of *E. coli*

Medium	Ingredients	Preparation and storage
LB medium	Bacto-pepton 10 g Yeast extract 5 g NaCl 10 g Distilled water up to 1000 ml	Adjust pH to 7.0, autoclave
SOC medium	Bacto-pepton 2 g Yeast extract 0,5 g 1 M NaCl 1 ml 1 M KCl 0,25 ml dissolve in 97 ml distilled water, autoclave and cool, then add: 2 M Mg^{2+} solution 1 ml 2 M Glucose solution 1 ml Distilled water up to 100 ml Store at -20 °C	
2 M Mg^{2+} solution	$MgCl_2.6H_2O$ 20,33 g $MgSO_4.7H_2O$ 24,65 g Distilled water up to 100 ml	Sterile Filtration
Glycerol freezing solution	80% Glycerol 10 mM $MgCl_2$	

2.6.2 Solutions for the Transformation and Selection of *E. coli*

Tab. 2-9: Solutions for the Transformation and Selection of *E. coli*

Name	Ingredients	Preparation and storage
100 mM IPTG solution	IPTG 0.24 g Distilled water up to 10 ml	Sterile filtration and storage at -20 °C
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactosid 100 mg N,N'-Dimethylformamid 2 ml	Wrap in aluminum foil and store at -20°C
Ampicillin	Ampicillin 50 mg Distilled water up to 1 ml	Sterile filtration and storage at -20 °C
Chloramphenicol	Chloramphenicol 50 mg Ethanol 100% up to 1 ml	Store at -20 °C

IPTG:

To prepare 10 ml of a 100 mM solution, 0.24 g IPTG was dissolved in sterile deionized water to give a final volume of 10 ml which was filter-sterilized through a 0.22 µm filter.

X-Gal:

For blue/white colony selection, 80 µg/ml of the artificial chromogenic substrate X-Gal were added to the agar plate just before use from a stock solution that was prepared as mentioned above (see 1.6.2).

Ampicillin:

A stock solution of 50 mg/ml was prepared in deionised water, filter-sterilized through a 0.22 µm filter, and stored at -20°C. To prepare selective medium, the medium was cooled to ~50°C after autoclaving and 1 ml of the ampicillin stock solution was added per liter of medium (both liquid and solid) to give a final concentration of 50 µg/ml.

Chloramphenicol:

A stock solution of 35 mg/ml was prepared in 100% ethanol. Filter-sterilization was not necessary. The stock solution was stored at -20°C. To prepare selective medium, the medium was cooled to ~50°C after autoclaving and 1 ml of the stock solution was added per liter of medium to give a final concentration of 35 µg/ml.

2.6.3 Buffer for Cell Lysis

Tab. 2-10: Buffer for Cell Lysis

Name	Ingredients	Preparation and storage
Potassium dihydrogen phosphate.	Potassium dihydrogenphosphate 1.36 g Water up to 100 ml	Adjust pH to 7.0 with KOH and store at 4°C.

2.6.4 Buffers and Solutions for DNA Electrophoresis

Tab. 2-11: Buffers and Solutions for DNA Electrophoresis

Name	Ingredients		Preparation and storage
50x TAE	Tris EDTA	2 M 0,05 M	Adjust pH to 8.0 with glacial acetic acid.
TBE buffer for sequencing Gel	Tris Boric acid EDTA	90 mM 90 mM 2 mM	
Loading buffer for sequencing	25 mM EDTA Formamide Dextran blue	1 ml 5 ml 0,1 mg	Formamide was dried with resin.
Sequencing gel	Urea 10x TBE Buffer Bidist. Water 30 % Acrylamid TEMED 10 % (w/v) APS	9 g 3,0 ml 11,5 ml 3,75 ml 0,01 ml 230 µl	
DNA loading buffer	Xylencyanol Bromphenol blue Ficoll 400 EDTA H_20	0.25 g 0.25 g 25.00 g 1.46 g up to 100 ml	The solutions were filtered and sterilized by autoclaving.

2.7 Enzymes for Molecular Biology

Tab. 2-12: Enzymes for Molecular Biology

Name	Specification	Supplier
DNA synthesis	Taq DNA polymerase Pwo DNA polymerase Pfu DNA polymerase Pfx DNA polymerase Superscript II Superscript III Thermoscript MMLV Reverse transcriptase	Life Technologies, MBI-Biotech, Peqlab Peqlab Stratagene Life Technologies Life Technologies Life Technologies Sigma, MBI-Biotech
Restriction enzymes	*Eco*RI, *Nhe*I.	Life Technologies
Other enzymes	T4 Ligase, DNase, RNase A, RNase H, and RNase T. Terminal transferase	Fermentas
DNA digestion	Calf Intestinal Phosphatase (CIP)	Ambion
RNA decapping	Tobacco Acid Pyrophosphatase (TAP)	Ambion

2.8 Primers

2.8.1 SMART-RACE and Vector Primers

Tab. 2-13: SMART-RACE and Vector Primers

Primer Name	Nucleotide Sequences 5`→ 3`
SMART II	5`- AAG CAG TGG TAA CAA CGC AGA GTA CGC GGG-3`
5'CDS	5`- $(T)_{25}$ $N_{-1}N$ -3`
3'CDS	5`- AAG CAG TGG TAA CAA CGC AGA GTAC$(T)_{30}$ $N_{-1}N$ -3`
RACE-short	5`- CTA ATA CGA CTC ACT ATA GGG C -3`
RACE-long	5`- CTA ATA CGA CTC ACT ATA GGG CAA GCA GTG GTA ACA ACG CAG AGT -3`
AAP	5`- GGC CAC GCG TCG AGT ACG GGI IGG GII GGG IIG-3`
AUAP	5`- GGC CAC GCG TCG ACT AGTAC -3`
M13 universal	5`- ACG ACG TTG TAA AAC GAG GGC CAG -3`
M13 reverse	5`- TTC ACA CAG GAA ACA GCT ATG ACC -3`
T7	5`- GAA TTG TAA TAC GAC TCA CTA TAG -3`
SP6	5`- GAT TTA GGT GAC ACT ATA GAA TAC -3`
3`pRSET reverse	5` –TAG TTA TTG CTC AGC GGT GG- 3`
RACE-nested	5`- AAG CAG TGG TAA CAA CGC AGA GT -3`

2.8.2 RLM-RACE Primers

Tab. 2-14: RLM-RACE Primers

Primer Name	Nucleotide Sequences 5`→ 3`
3`RACE outer	5`- GCG AGC ACA GAA TTA ATA CGA CT -3`
3`RACE inner	5`- CGC GGA TCC GAA TTA ATA CGA CTC ACT ATA GG -3`
5`RACE outer	5`- GCT GAT GGC GAT GAA TGA ACA CTG -3`
5`RACE inner	5`- CGC GGA TCC GAA CAC TGC GTT TGC TGG CTT TGA TG-3`
3`RACE adapter	5`-GCG AGC ACAGAA TTAATA CGA CTC ACT ATA GGT $_{12}$VN-3`
3`RACE control	5`- AGC AGT TGG TTG GAG CAA ACA TC -3`
5`RACE outer control	5`- GAT CAC CAA TCC ATT GCC GAC TAT -3`
5`RACE adapter	5`- GCU GAU GGC GAU GAA UGA ACA CUG CGU UUG CUG GCU UUG AUG AAA -3`
5`RACE inner control	5`- GAA GTA GAT GGT GGG CAG GAA GAT -3`
PCR control	5`- GCA GCA GGT AGC AGT GAC -3`

2.8.3 Degenerate Primers for Use in *H. androsemum* and *S. acuoparia*

Tab. 2-15: Degenerate Primers for Use in *H. androsemum* and *S. acuoparia*

Primer Name	Nucleotide Sequences 5`→ 3`
CoA-ligase reverse	5`- TT(AG) CC5 AT(CT) TC5 TC5 GG5 AC5 AC -3`
CoA- ligase forward1	5`- C5G C(CT)T C5G TCA T5C C(AG)T A5C C(CT)T G -3`
CoA- ligase forward2	5`- CC5 C(GT) (ATG) AT(AG) CA(AGT) AT(CT) TC5 CC -3`

2.8.4 Gene Specific Primers (GSP) for Use in *H. androsaemum*

Tab. 2-16: Gene Specific Primers (GSP) for Use in *H. androsaemum*

Primer Name	Nucleotide Sequences 5`→ 3`
Upper forward	5`- CGT ATT TCG TCA TGG GGC TGT TG -3`
Lower reverse	5`- TTC CAC AAT TTT TCG CAC TCC TGA -3`
Frg. 1 forward	5`- CAG GGC TAT GGC ATG ACC GAG GCC -3`
Frg. 2 forward	5`- CTA TTC TCG TGT GTC AAG ATC ACA C -3`
Frg. 3 forward	5`- CCG GAT GAC ACC GTG CAA ACC -3`
GSP2 forward	5`- AAA GGT GTG ATC TTG AGA CAG -3`
GSP2 reverse	5`- CGC ACC AGT AGT ACC CGA AGA -3`

2.8.5 Gene Specific Primers for CoA-ligases from *S. aucuparia*

Tab. 2-17: Gene Specific Primers for CoA-ligases from *S. aucuparia*

Primer Name	Nucleotide Sequences 5`→ 3`
Sorbus reverse1	5`- CTA TAA CCA GCG GCG GCA CCA -3`
Sorbus reverse2	5`- CAA CGG CAG CAC GCA CAA TAC GAC -3`
Sorbus sequence	5`- GAG CTC TAT TAG GGT GGT GTT GTC -3`
Sorbus 1 A forward	5`- GGC TAA GTT ACC TAG TGC CAA AC -3`
Sorbus 1 B forward	5`- TTT GTA CCT CCG ATC GTT TTG AGT -3`
Short-1 Sorbus reverse	5`- GAG TCC GCA GAG AAA CAC TGA A -3`
Short-2 Sorbus reverse	5`- GTT TGG CAC TAG GTA ACT TAG -3`
Missing-1 Sorbus reverse	5`-GTA GGT GTG GAG AGG GAG GTG GTT -3`
Missing-1 Sorbus reverse	5`-GTA GAA GGG GTT GGC GGT GGT -3`
Over . Exp. 3’	5`-GAA TTC TTA GGG CAG CGG TGG GGT TGC GGTG-3`
Over. Exp. 5’	5`- ATG ATC TCC ATT GCT TCT AAT TCC GTT -3`
Seq . 1 new reverse	5`- AAT CGT CAC CTT GTA ATT CTC CAC C -3`
Sorbus2 A forward	5`- TTCAAC AGC GTG CTG GCT GTG -3`
Sorbus2 B forward	5`- GGT GAG CTT CTG GAG CTC ATT -3`
Sorbus3 A forward	5`- ACA TGC TGG CTG CTC TAG CCG CA -3`
Sorbus3 B forward	5`- GAG ACC TCT GGG AAA AGA TCC GG -3`
Sorbus4 A forward	5`- GTC GCA CAA AAT CGC GCT AAA CAG C -3`

Sorbus4 B forward	5`- GTG TAT TAC GAG AGC TGC TCC TC -3`
Short-2 Sorbus reverse	5`- GTT TGG CAC TAG GTA ACT TAG -3`
Seq . 1 new reverse	5`- AAT CGT CAC CTT GTA ATT CTC CAC C -3`
Seq. 2 reverse	5`-CCGACA CTC GGT ACC GCT GAA TG -3`
Seq. 2 new reverse	5`-AAC ACT GCC TGA GGG ACT CGG CTC TTC A -3`
Seq. 3 reverse	5`- GGT ATT TAT GAG ATT TTC CGG A -3`
Seq. 3 new reverse	5`- TAC CCT GTC CCA AGT TTG GCA -3`
Seq. 4 reverse	5`- CTT GTT GGT CCC AAA GAG GAA G -3`
Seq. 4 new reverse	5`- CAG CCA GCT TTG CTC TCA AGT CTT TTC -3`
15 GSP reverse	5`- GAC TAG GCT CGT CCG ATA AAG CGA A-3`
15 GSP new reverse	5`- CTT GAT GAC GAT AAG TCC TTG ACC -3`
15 GSP forward	5`- CGC TTA CGC GGA CGG TTG CGT AGC -3`
5´- 15 GSP reverse	5`- GGT TTA AAG ATC TCG GCC GGT -3`

2.9 Molecular Biology Materials

2.9.1 Kits

2.9.1.1 Nucleic Acids Isolation and Purification Kits

Tab. 2-18: Nucleic Acids Isolation and Purification Kits

Name	Specification	Supplier
RNA isolation	QuickPrep® micro mRNA Purification Kit	Amersham Biosciences
	Rneasy® Plant mini Kit	Qiagen
	Neuclospin® RNA plant	Macherey –Nagel
DNA isolation and purification	Concert™ Rapid Plasmid Miniprep System	Life Technologies
	Concert™ Rapid Gel Extraktion System	Life Technologies
	Mini Elute Gel Extraktion Kit	Qiagen
	Nucleospin®E.Z.N.A.Plasmid Miniprep Kit	Peqlab
	Neuclospin® Extract II Gel Extraction kit	Macherey-Nagel
	Neuclospin® Plasmid Kit	Macherey-Nagel

2.9.1.2 Amplification of 3`and 5`Ends of cDNA

Tab. 2-19: Amplification of 3`and 5`Ends of cDNA

Name	Supplier
5` RACE (Rapid amplification of cDNA ends) kit	Invitrogen, and Clontech
RNA Ligase mediated rapid amplification of cDNA Ends (RLM-RACE-PCR) kit	Ambion

2.9.1.3 Cloning of DNA

Tab. 2-20: Cloning of DNA

Name	Supplier
pGEM easy T vector	Invitrogen
pRSET-B vector (Expression vector)	Invitrogen

2.9.2 Buffers and Solutions for Plasmid DNA Isolation (mini-prep.)

Tab. 2-21: Buffers and Solutions for Plasmid DNA Isolation

Name	Ingredients		Preparation and storage
Buffer P1	Tris-HCl EDTA RNase A	50 mM 10 mM 100 µg/ml	Tris-HCl and EDTA solutions were filter-sterilized and mixed. The pH was adjusted to 8.0. RNase A was freshly added before use.
Buffer P2	NaOH SDS	0,2 M 1 % (w/v)	Autoclaved.
Buffer P3	Potassium acetate	2,55 M	Prior to autoclaving, the pH was adjusted to 4.8.
Sodium acetate solution	Sodium acetate	3 M	The pH was adjusted to 4.0.
TE-buffer	Tris-HCl EDTA	10 mM 1 mM	The pH was adjusted to 7.0

2.9.3 Host Cells

Tab. 2-22: Host Cells

E.coli	Genotype
K12® DH 5α	*F' φ80δlacZΔM15 end A1 hsdR17(rk^-mk^+) supE44 thi-1 λ^- gyrA96 relA1 Δ(lacZYA-argFV169) deoR*
JM 109	*e14$^-$(McrA$^-$) recA1 endA gyrA96 th-1 hsdR17(rk^-mk^+) supE44 relA1 Δ(lac-proAB) [F' traD36 proAB lac^qZ ΔM15*
NM 522	*hsd⁄Δ5⁄Δ(lac-pro) [F' lacIqZ⁄Δ M15 pro^+]*
BL 21	*F^-, ompT, hsdS (r_B^-, m_B^-), gal, dcm*
BL21(DE3)pLysS	*F^-, ompT hsdSB (r_B^- m_B^-) gal dcm (DE3) pLysS (CamR)*

3. METHODS

3.1 Nucleic Acids

3.1.1 Isolation of mRNA

mRNA was isolated using the QuickPrep micro mRNA Purification Kit (Amersham Biosciences) based on the combined disruptive and protective properties of guanidinium thiocyanate (GTC) with the speed and selectivity of oligo(dT)-cellulose chromatography in a spun-column. Cultured cells (100 mg; dry weight) were homogenized under liquid nitrogen in a mortar with seasand. Initially, the cells were extracted by homogenization in a buffered solution containing a high concentration of GTC. This ensures the rapid inactivation of endogenous RNase activity and the complete dissociation of cellular components from the mRNA. The extract was then diluted three-fold with Elution Buffer to reduce the GTC concentration to a level low enough to allow efficient hydrogen-bonding between poly(A) tracts on mRNA molecules and oligo(dT) attached to cellulose, but high enough to maintain complete inhibition of RNases. As an additional benefit, the three-fold dilution causes a number of proteins to precipitate, allowing them to be removed easily by centrifugation. After a brief second homogenization, the extract was clarified by centrifugation, the supernatant was poured into an Oligo(dT)-Cellulose Spun Column, and the polyadenylated fraction was allowed to bind over a short period of time with gentle mixing. The column was subjected to a low speed centrifugation, and the liquid containing the non-bound material was decanted. The matrix was batch-washed sequentially with High-Salt and Low-Salt Buffer. Finally, the sample was eluted from the matrix with prewarmed Elution Buffer. High-quality, full-length mRNA is necessary for meaningful results in any gene expression study. Therefore, choosing the optimal mRNA sample preparation method was essential, and all material that came into contact with the sample had been treated to prevent RNase contamination, and all reagents used were RNase-free.

3.1.2 Isolation of Total RNA

The isolation and purification of total RNA from cell lysates was performed using the RNeasy® mini kit (Qiagen). According to the manufacturer's instructions, 0.1 g cells were first ground in liquid nitrogen and then lysed under highly denaturing conditions. 10 µl of the provided ß-mercaptoethanol were added per 1 ml of RLT lysis buffer. Cells were resuspended well in 450 µl RLT buffer and frozen at -80°C. Frozen lysates were thawed at 37°C for 10 min and transferred directly into QIAsheder columns sitting in 2 ml collection tubes and centrifuged at maximum speed for 2 min. Samples were collected into new tubes and 0.5 volume 250 µl of ethanol (96-100%) was added, creating conditions which promote selective binding of RNA to the RNeasy® matrix. 700 µl of sample were applied to a RNeasy mini spin column sitting in 2 ml collection tubes and centrifuged at 10,000 rpm for 15 sec to allow absorption of RNA to the matrix. If the volume of lysate was exceeding 700 µl, all the rest of suspension was applied to the same RNeasy® spin column using the same collection tube but discarding each time the flow-through. The columns were washed with 700 µl of RW buffer and centrifuged at 10,000 rpm for 15 sec. After having transferred the RNeasy® spin columns into new 2 ml collection tubes, 500 µl of RPE buffer containing ethanol were added and centrifuged at 10,000 rpm for 15 sec. columns were recentrifuged at maximum speed for 2 min to dry the RNeasy® membrane. The RNeasy® spin columns were transferred into 1.5 ml collection tubes and RNA was eluted with two centrifugation steps at 10,000 rpm for 1 min with 40 µl of sterile distilled water. RNA was stored at -80°C. All the following procedures were performed keeping RNA samples on ice.

3.1.3 Isolation of DNA from Phages

A bacteriophage is a virus that infects bacteria and can either instantly kill a bacterial cell or integrate its DNA into the host bacterial chromosome (Madigan et al., 1997). phages have a dual existence. They can undertake the lytic pathway, or integrate their chromosome into the host cell chromosome and replicate every time the host replicates, this latent form of the phage is called a prophage, and this form of propagation is called lysogeny. The phage lysate contains phage DNA encased in the phage coat. Phages capable of lysogeny and lytic pathways are able to switch between the two processes. In the lytic pathway, viral functions are fully expressed: viral DNA and proteins are quickly produced and packaged into virus particles, leading to the lysis (destruction) of the host cell and the sudden appearance of about

100 progeny virus particles, or virions. A cDNA library complementary to *H. androsaemum* mRNA was available in the laboratory. Preparation of phage DNA included the following steps: digestion of bacterial DNA, precipitation of phage particles, removal of the phage coat by organic extraction, and purification of phage DNA by precipitation with alcohol. RNaseA and DNaseI were added to a final concentration of 1µg/ml each. Phage lysate (5 ml) was incubated in a 30 ml Falcon tube with the nuclease at 37°C for 30 minutes. For precipitation, 4 ml 20% PEG (polyethylene glycol) and SM buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM MgSO4, 0.01% gelatin) were added and mixed well. The mixture was incubated on ice for 1 hr and then centrifuged at 10.000 rpm for 10 min at 4°C. The supernatant was discarded and the pellet was dissolved in 0.6 ml of SM buffer and transferred to a new 1.5 ml tube. 2.5 ml 20 % SDS and 5 ml 0.5 M EDTA were added and the mixture was incubated for 15 min in a water bath at 68°C. The mixture was extracted twice with equal volumes of phenol: chloroform (1:1 v/v). After centrifugation at 14.000 rpm for 10 min at room temperature, the upper aqueous layer was transferred to a new 1.5 ml tube. The next steps were addition of isopropyl alcohol, incubation at -80°C for 12 min, and centrifugation at 14.000 rpm for 20 min at room temperature. To wash the pellet, 0.7 ml of ethanol (70%) was added and centrifuged at 14.000 rpm for 10 min. The supernatant was discarded, and the pellet was dried in the vacuum for 10 min, dissolved in 50 µl of TE buffer, and stored at -20°C.

3.2 cDNA Synthesis

3.2.1 5` and 3`- RACE using the RACE-SMART method

RACE is a PCR-based method for generating high yields of high-quality, double-stranded cDNA from total or poly(A)+ RNA. SMART stands for **S**witching **M**echanism **A**t 5' end of **R**NA **T**emplate. cDNA is a strand of DNA complementary to appropriate mRNA. First strand cDNA was synthesized from mRNA in a reaction catalyzed by reverse transcriptase, an RNA dependent-DNA polymerase. The synthesis of cDNA was carried out as following:

Tab. 3-1: 3`and 5`-RACE cDNA synthesis

3`- RACE-cDNA		**5`-RACE -cDNA**	
mRNA	3-5 µl	mRNA	3-5 µl
3`-CDS primer	1 µl	Smart II oligo primer	2 µl
Nuclease-free water up to	10 µl	5`-CDS primer	1 µl
		Nuclease-free water up to	10 µl

The components were mixed and briefly centrifuged. The tubes were incubated at 70°C for 2 min. The PCR tubes were cooled on ice for 2 min and briefly centrifuged to collect the contents at the bottom. The following components then were added to each reaction tube (already containing 5 µl):

5x First-strand buffer	2 µl
DTT (20 mM)	1 µl
200 U /µl Reverse transcriptase	1.25 µl
dNTP mix (10 mM)	1 µl

The contents of the tubes were mixed gently by pipetting and centrifuged briefly to collect them at the bottom. The reaction mixture was incubated at 42^{o}C for 60 min and then heated, for inactivation of the enzyme, at 70^{o}C for 15 min. The cDNA sample was either used for PCR or stored at -20^{o}C. All primers used in this procedure are listed under (2.8.1).

3.2.2 cDNA Synthesis using Terminal Transferase

The Terminal Deoxynucleotidyl Transferase (TdT), a template-independent DNA polymerase, catalyzes the repetitive addition of deoxyribonucleotides to the 3'-OH of oligodeoxyribonucleotide and single-stranded, and double-stranded DNA. The TdT requires an oligonucleotide of at least three nucleotides to serve as a primer.

The following reaction mixture was performed:

5X reaction buffer		4 µl
DNA fragment		1 µl 3`-end
dNTP		1 µl
Terminal Deoxynucleotidyl Transferase		20-40 U
Water, nuclease free	up to	20 µl

3.2.3 cDNA Synthesis using RLM-RACE

RNA **L**igase **M**ediated **R**apid **A**mplification of **c**DNA **E**nds (RLM-RACE) represents a major improvement to the classic RACE technique (Maruyama and Sugano, 1994; Shaefer, 1995). RLM-RACE is designed to amplify cDNA only from full-length, capped mRNA, usually producing a single band after PCR. Total or poly(A) selected RNA was treated with calf intestinal phosphatase (CIP) to remove free 5' phosphates from molecules such as ribosomal RNA, fragmented mRNA, and tRNA. The cap structure found on intact 5' ends of mRNA were not affected by CIP. The RNA was treated with Tobacco Acid Pyrophosphatase (TAP) to remove the cap structure from full-length mRNA, leaving a 5' monophosphate. A synthetic RNA adapter was ligated to this RNA population using T4 RNA ligase. The adapter cannot ligate to dephosphorylated RNA because these molecules lack the 5' phosphate necessary for ligation. During the ligation reaction the majority of full-length decapped RNA acquires the adapter sequence as the 5' end. A random-primed reverse transcription reaction and nested PCR were then performed to amplify the 5' end of a specific transcript. Two nested primer corresponded to the adapter sequence and the nested antisense primer was specific for the target gene. RLM-RACE permits ligation of the synthetic RNA adapter only to decapped (full-length) RNA.

Calf Intestinal Phosphatase (CIP) treatment:

Total RNA		1 µg
10x CIP buffer		2 µl
Calf Intestinal Phosphatase (CIP)		2 µl
Nuclease-free Water	up to	20 µl

The samples were gently mixed, briefly centrifuged, and incubated at 37°C for 1h.

Tobacco Acid Pyrophosphatase (TAP) treatment:

CIP treated RNA (from above)	4 µl
10x TAP buffer	1 µl
Tobacco acid pyrophosphatase	1 µl
Nuclease-free water	2 µl

The samples were gently mixed, briefly centrifuged, and incubated at 37°C for 1h. Reaction product was used for the ligation step.

RNA Adapter Ligation:

CIP/TAP-treated RNA	5 µl
RNA adapter	1 µl
10x RNA ligase buffer	1 µl
T4 RNA ligase (2.5 U/µl)	2 µl
Nuclease-free water	5 µl

The samples were gently mixed, briefly centrifuged, and incubated at 37°C for 1h. Reaction product was used for reverse transcription.

Reverse Transcription:

Ligated RNA		2 µl
dNTP mix		4 µl
Random decamers or gene specific primer		2 µl
10x RT-PCR buffer		2 µl
RNase inhibitor		1 µl
MMLV reverse transcriptase		1 µl
Nuclease-free water	up to	20 µl

The samples were gently mixed, briefly centrifuged, and incubated at 42°C for 1 h. Reaction product was used for the PCR step. All enzymes and primers used in this procedure are listed under (2.7 and 2.8.2).

3.3 Phenol/Chloroform Extraction in Nucleic Acids Preparation

For extraction of nucleic acids, one volume of phenol/chloroform solution was added. After mixing and centrifugation at 15.000 rpm for 5 min at 4°C, the upper phase was removed and the phenol/chloroform step repeated twice. At the end of the extraction, the upper layer was used for ethanol precipitation.

3.4 Precipitation of DNA

An ethanol / 7 M ammonium acetate solution (2.5-3 volumes) was added to the DNA sample in a microcentrifuge tube, and the mixture was centrifuged at 14.000 rpm for 20 min at 4°C. To recover the precipitated DNA, the supernatant was discarded and the DNA pellet washed with a 70% ethanol solution. After a second centrifugation, the supernatant was again discarded. Pellet was dried at clean bench, and then dissolved with 30-50 µl distilled water. The DNA solution was preserved at -20°C.

3.5 Purification of DNA

Different purification kits were used to obtain pure DNA. Following gel electrophoresis, DNA fragments were eluted from the gel as described under (3.12.2). DNA was adsorbed to a silica gel membrane in the presence of a high salt concentration. Contaminants were washed away using wash buffer, and the DNA was eluted in 20 µl TE buffer or distilled water.

3.6 Determination of Nucleic Acids Concentration

3.6.1 Spectrophotometric Determination

DNA concentration was determined by measurement of the adsorption at 260 nm. An Ultraspec1000 Pharmacia Biotech spectrophotometer was used to determine the concentration of 1:50 diluted RNA or DNA samples in a volume of 0.5 ml in a 10 mm Light Patch Quarzcuvette. Absorbance readings (*A*260) should fall between 0.1 and 1.0 to be accurate. Sample dilution should be adjusted accordingly. An absorbance of 1.0 at 260 nm corresponds to 50 µg DNA/ml. Since spectrophotometric measurement cannot differentiate between PCR products, primers, and oligonucleotides, the PCR product should be purified prior to measurement.

3.6.2 Estimation by Agarose Gel Analysis

The concentration of DNA or RNA solutions was roughly quantitized by agarose gel analysis. The sample was applied to an agarose gel side-by-side with known amounts of a marker of

the same size. The DNA or RNA can be visualized under UV light following staining with ethidium bromide. The concentration of the sample was estimated by visual comparison of the intensity of the sample band with that of the standard bands.

3.7 Polymerase Chain Reaction (PCR)

The polymerase chain reaction serves to amplify DNA, with traces of DNA being sufficient. It uses repeated cycles, each of which consists of three steps:

1. The reaction solution containing DNA molecules (to be copied), polymerase (to copy the DNA), primers (to serve as starters) and nucleotides (to extend the primers) is heated to 95°C. This causes the two complementary strands to separate, a process known as denaturing or melting.
2. Lowering the temperature to 55°C causes the primers to bind to the DNA, a process known as hybridization or annealing. The resulting hydrogen bonds are stable only if the primer and the DNA segment are complementary, i.e. if the base pairs of the primer and the DNA segment match. The polymerase then begins to attach additional complementary nucleotides at these sites, thus strengthening the bonding between the primers and the DNA.
3. Extension: The temperature is again increased, this time to 72°C. This is the ideal working temperature for the polymerase used, which adds further nucleotides to the growing DNA strand. At the same time, any loose bonds that have formed between the primers and DNA segments that are not fully complementary are broken. Each time these three steps are repeated the number of copied DNA molecules doubles. After 20 cycles about a million molecules are amplified from a single segment of double-stranded DNA. The temperatures and duration of the individual steps described above refer to the most commonly used protocol. A number of modifications have been introduced that give better results to meet specific requirements.

3.7.1 Hot start PCR

PCR primers and templates match more specifically at higher temperatures, but hot reagents can be hard to work with. That’s why hot start PCR uses a variety of barriers to keep the reaction components separate until the heat is on. In hot-start PCR, polymerase activity was

prevented during reaction setup and starts only after the initial denaturation step. Hot-start technology generates cleaner PCR products. The methodology prevents non-specific extension or degradation of nucleic acid substrates at ambient temperatures by either excluding or reversibly inhibiting the polymerase enzyme. Upon assembly, pre-heating the other reaction components melts all priming events, both specific and non-specific. Addition of the polymerase, if missing, then initiates PCR. Alternatively, the heat also reverses the inhibition of the enzyme thus activating it. The first annealing step, due to its properly defined temperature, allows specific annealing reactions to occur and prevents non-specific annealing events. With a lack of non-specific hybridization of primers to template or to one another, the resulting amplified DNA bands are cleaner.

3.7.2 Reverse Transcriptase - PCR (RT-PCR)

RT-PCR (Reverse Transcription-Polymerase Chain Reaction) is a technique in which an RNA strand is reversely transcribed into its DNA complement, followed by amplification of the resulting DNA using a polymerase chain reaction (PCR). Transcribing an RNA strand into its DNA complement is termed reverse transcription (RT) and is accomplished through the use of an RNA-dependent DNA polymerase (reverse transcriptase). The RNA template is degraded by RNaseH. A second strand of DNA is synthesized through the use of a deoxyoligonucloteide primer and a DNA-dependent DNA polymerase. The complementary DNA and its anti-sense counterpart are exponentially amplified via a polymerase chain reaction (PCR).

3.7.3 Touch-down PCR

Touch-down PCR uses a cycling program with varying annealing temperatures. It is a useful method to increase the specificity of PCR. The annealing temperature in the initial PCR cycle should be 5–10°C above the estimated melting temperature(*T*m) of the primers. In subsequent cycles, the annealing temperature is decreased in increments of 1–2°C per cycle until a temperature was reached equal to, or 2–5°C below, the *T*m of the primers. Touchdown PCR enhances the specificity of the initial primer–template duplex formation and hence the specificity of the final PCR product.

3.7.4 Nested PCR

If PCR sensitivity is low, nested PCR often improves product yield. This technique involves two rounds of amplification reactions. The first-round PCR is performed according to the PCR protocol using *Taq* DNA polymerase. Subsequently, an aliquot of the first-round PCR product was subjected to a second round of PCR. The second-round PCR was performed with two new primers that hybridize to sequences internal to the first-round primer target sequences. In this way, only specific first-round PCR products (and not non-specific ones) will be amplified in the second round. Alternatively, it is possible to use one internal primer and one first-round primer in the second PCR, which is referred to as semi-nested PCR.

3.7.5 Rapid Amplification of cDNA Ends (RACE) PCR

PCR can be used to facilitate isolation of 3`and 5'-ends of mRNA by several similar methods collectively termed Rapid Amplification of cDNA Ends, or RACE. RACE involves performing a randomly-primed reverse transcription (RT) reaction, adding an adapter to the 3' and 5`-end of the synthesized cDNA by either ligation or polymerase extension, and amplifying by PCR with a gene specific primer and a primer that recognizes the adapter sequence. 3’ and 5’ RACE methodologies offer possible solutions to this problem. 3’ RACE takes advantage of the natural poly (A) tail in mRNA as a generic priming site for PCR amplification. In this procedure, mRNAs were converted into cDNA using reverse transcriptase (RT) and an oligo-dT adapter primer. Specific cDNA was then directly amplified by PCR using a gene-specific primer (GSP) that anneals to a region of known exon sequences and an adapter primer that targets the poly(A) tail region. This permits the capture of unknown 3’-mRNA sequences that lie between the exon and the poly(A) tail.

5’ RACE is a technique that facilitates the isolation and characterization of 5’ ends from low-copy messages. First strand cDNA synthesis is primed using a gene-specific antisense oligonucleotide. This permits cDNA conversion of specific mRNA, or related families of mRNAs, and maximizes the potential for complete extension to the 5’-end of the message.

3.8 Primer design

The most critical parameter for successful PCR is the design of primers. A poorly designed primer can result in an unsuccessful PCR reaction. The primer sequence determines several things such as the length of the product, its melting temperature and ultimately the yield. A badly designed primer can result in little or no product due to non-specific amplification and/or primer-dimer formation, which can become competitive enough to suppress product formation. This application note is provided to give rules that should be taken into account when designing primers for PCR (Dieffenbach et al., 1995). The sequences of the primers used for PCR amplification can have a major effect on the specificity and sensitivity of the reaction. When choosing two PCR amplification primers, the following guidelines should be considered:

1. Primers should be at least 18 nucleotides in length to minimize the chances of encountering problems with a secondary hybridization site on the vector or insert.

2. Primers with long runs of a single base should generally be avoided. It is especially important to avoid 4 or more G's or C's in a row.

3. For cycle sequencing, primers with melting temperatures in the range 52-58°C generally produce better results than primers with lower melting temperatures. Primers with melting temperatures above 65°C should also be avoided because of potential for secondary annealing. If you are working with high “G-C” templates, then a primer with a Tm 60°C or above may be desirable. It is then advisable to do the sequencing reaction with annealing and extension at 60°C.

4. Primers should have a G/C content between 40 and 60 percent. For primers with a G/C content of less than 50%, it may be necessary to extend the primer sequence beyond 18 bases to keep the melting temperature above the recommended lower limit of 50°C.

5. Primers should be "stickier" on their 5' ends than on their 3' ends. A "sticky" 3' end as indicated by a high G/C content could potentially anneal at multiple sites on the template DNA. A "G" or "C" is desirable at the 3' end but the first part of this rule should apply.

6. Primers should not contain complementary parts (palindromes) within themselves; that is, they should not form hairpins. If this state exists, a primer will fold back on itself and result

in an unproductive priming event that decreases the overall signal obtained. Hairpins that form below 50°C generally are not such a problem.

7. Primers should not contain sequences of nucleotides that would allow one primer molecule to anneal to itself or to the other primer used in PCR reactions (primer-dimer formation).

3.8.1 Degenerate Primers

A PCR primer sequence is called degenerate if some of its positions have several possible bases. The degeneracy of the primer is the number of unique sequence combinations it contains. Degeneracy in primer sequence should also be taken into consideration. Degenerate primers based on the amino acid sequence of conserved regions were used to search for members of a gene family (Wilks et al., 1989). The use of degenerate PCR has proven to be a very powerful tool to find new genes or gene families. Most genes belong to families that share structural similarities. By aligning the protein sequences from a number of related proteins, one can find shared conserved sequences which can be used as a starting point to make degenerate PCR primers. Several general rules for the construction of degenerate pools of primers from conserved amino acid motifs are as follows:

1. Use codon bias of the appropriate species for translation.
2. Use two blocks of conserved amino acids for primer pairs.
3. Use 4-6 amino acid sequences as bases for degenerate primers. The length of the primers should be a minimum of 20 nucleotides.
4. Avoid degenerate bases at 3' end, omit the last base of terminal codon unless the amino acid is met or trp.
5. Consider using deoxyinosine to reduce degeneracy.
6. Primers ending in T are efficiently extended even when mismatched with T, G or C. 3' terminal mismatches A-G, G-A, C-C, and G-G reduce PCR yields 100 fold and A-A mismatch reduces yield by 20 fold.

3.8.2 Design of Gene Specific Primers (GSP)

Efficient and specific PCR is highly dependent on effective primer design. This is especially true for RACE applications since the PCR is carried out with only a single GSP. No method

of primer design can guarantee successful amplification, so all primers must be tested in PCR before they can be pronounced good. If introduction of a restriction site is required, it should be flanked by 3-10 (depending on the restriction enzyme "spacer") nucleotides at the end of the primer to allow an efficient digestion close to the end of the cDNA. The primers used in this work (2.8.3, 2.8.4, 2.8.5) were synthesized by the MWG-Biotech company.

3.9 Cloning of PCR Products

All new recombinant plasmids generated in this work were prepared by ligating into the appropriate recipient vector the gene of interest amplified by PCR and containing the compatible cloning restriction sites. Briefly, the strategy was as follows:

- Intend and synthesize 5' and 3'-end PCR primers with the desired restriction sites.
- Perform the PCR with the appropriate primers and DNA template
- Clone the PCR product into the pGEM T Easy Vector
- Transformation of DNA into DH5α^{TM}
- Analysis of transformants (blue/white selection)
- Plasmid DNA minipreparations
- Identification of clones containing the correct DNA insert (restriction analysis)
- Insert DNA sequence analysis
- Preparation of the pRSETB vector and the insert for cloning
- Ligation of plasmid vector and insert DNAs
- Transformation of DNA into BL21 plasmid DNA minipreps
- Identification of clones containing the correct DNA insert (restriction analysis)
- Insert DNA sequence analysis
- Store the correct plasmid DNA and bacteria as glycerol stock at -80°C

3.9.1 Cloning in pGEM-T Easy Vector

The pGEM®-T Easy Vector was used to clone target DNA. The vector contains T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region of the enzyme β-galactosidase. Insertional inactivation of the α-peptide allows recombinant clones to be directly identified by color screening on indicator plates. The multiple cloning regions of the vector include restriction sites which allow for the release of

the insert by digestion with a single restriction enzyme. The multiple cloning region is flanked by recognition sites for the restriction enzymes EcoRI, BstZ I and NotI, thus providing three single-enzyme digestions for release of the insert. Alternatively, a double-digestion may be used to release the insert from either vector. The vector was prepared by cutting with EcoRI and adding a 3´-terminal thymidine to both ends. These single 3´-T overhangs at the insertion site greatly improve the efficiency of ligation of a PCR product into the plasmids by preventing recircularization of the vector and providing a compatible overhang for PCR products generated by certain thermostable polymerases. These polymerases often add a single deoxyadenosine, in a template-independent fashion, to the 3´-ends of the amplified fragments.

3.9.2 Cloning in pRSET B (Invitrogen) Expression Vector

The pRSET B vectors are pUC-derived expression vectors designed for high-level protein expression and purification from cloned genes in *E. coli*. DNA inserts were positioned downstream and in frame with a sequence that encodes an N-terminal fusion peptide. This sequence includes an ATG translation initiation codon.

3.9.3 Ligation of Plasmid Vector and DNA Insert

After the vector and insert DNAs have been purified, the concentration of DNA was estimated by agarose gel electrophoresis along with molecular weight standards. The ligation reaction was performed using the T4 DNA Ligase according to the manufacturer's instructions. The typical ligation reaction was performed at 4°C over night, with 50-200 ng of vector DNA (molar ratio of vector to insert 1:3) in 10 µl final volume:

Plasmid vector		1-2 µl
DNA insert		3-6 µl
10x DNA ligase buffer		2 µl
T4 DNA ligase		1 µl
Water	up to	10 µl

To monitor the efficiency of ligation of vector with insert, the linear vector was subjected to ligation in the absence of the DNA insert. At the end of the reaction, the mixture samples were placed on ice and 5µl were transformed into competent cells (3.10.2.2).

3.10 Host Cells and Transformation

The purpose of this technique is to introduce a foreign plasmid into bacteria and to use those bacteria to produce large quantities of the plasmid. This is based on the natural function of a plasmid to transfer genetic information vital to the survival of the bacteria under given conditions. In order to make bacteria take up a plasmid, they must first be made "competent" for this process. This was done by creating small holes in the bacterial cells by suspending them in a solution of high calcium concentration. DNA can then be forced into the cells by incubating them together with the DNA on ice, placing them briefly at 42^{o}C (heat shock), and putting them back on ice. This causes the bacteria to take up the DNA. The cells are finally plated out on an antibiotic-containing medium.

3.10.1 Host Cells

3.10.1.1 DH5α^{TM}

This strain of *E. coli* was used for both initial cloning of target DNA into pGEM T Easy vector and plasmid maintenance, since it is a *recA* and *endA* mutant that allows high transformation efficiency and high plasmid yield. The cells contain the coding information for the carboxy-terminal portion of ß-galactosidase (*lacZ*) that, in association with the corresponding vector, allows blue/white colony selection.

3.10.1.2 BL21 (DE3)

BL21 is the strain of choice for high-level gene expression and production of recombinant proteins in bacterial systems. BL21 strains lack the *lon* and *ompT* proteases thus promoting stability of recombinant proteins. Strains designated as “DE3” carry a copy of the T7 RNA

polymerase under control of the IPTG inducible *lacUV5* promoter, and as a result, are ideal for controlled expression of T7 promoter-driven constructs. This is required for transcription of the target gene cloned in pRSET-B vector.

3.10.1.3 BL21 (DE3) PLys S

This strain carries the DE3 bacteriophage lambda lysogen. This lambda lysogen contains the *lac*I gene, the T7 RNA polymerase gene under control of the *lacUV5* promoter, and a small portion of the *lac*Z gene. This *lac* construct is inserted into the *int* gene which is hereby inactivated. Disruption of the *int* gene prevents excision of the phage (i.e. lysis) in the absence of helper phage. The *lac*-repressor represses expression of T7 RNA polymerase. Addition of IPTG allows expression of this polymerase. Strains carry the pLysS plasmids encoding T7 lysozyme, a natural inhibitor of T7 RNA polymerase, which reduces background levels of polymerase activity in uninduced cells. When the recombinant protein is toxic to the cell, the pLysE host is ideal because it produces higher levels of the inhibitor providing the most stringent control in T7 RNA polymerase promoter based systems.

3.10.2 Transformation.

3.10.2.1 Preparation of competent cells

Competent cells were prepared according to a modified method developed by Cohen et al. (1972). Cells were grown in 50 ml liquid LB medium at 37°C with shaking at 225 rpm for 6-8 h until the bacterial culture reached an optical density of 0.5-1 at 600 nm. The cells then were chilled on ice for 10 min and centrifuged at 4°C and 3500 rpm for 10 min. The bacterial pellet was resuspended carefully in 20 ml ice-cooled glycerol containing 50 mM $CaCl_2$. The cell suspension was then kept at 4°C for 20-24 h to achieve high competency. Competent cells were stored at -80°C in 50 µl or 100 µl aliquots until use.

3.10.2.2 Transformation of Competent Cells

This procedure was performed under sterile conditions. Use only autoclaved plastic-ware and always work with a flame. Also, bacteria are very labile at high calcium, so keep the bacteria on ice and away from the flame at all times to keep them viable. For transformation, the plasmid DNA (10-100 ng) was added to a 50-100 µl aliquot of competent cells (3.10.2.1). After incubation on ice for 30 minutes, cells were heat-shocked at 42°C for 45 sec in a water bath. The tube was then placed on ice for 5 min. 700 µl of SOC media was added, and the culture was incubated for 1 h at 37°C with vigorous shaking (> 225 cycles/min in a rotary shaker), and centrifuged at 5000 rpm for 5 min. Appropriate amounts of cells were plated on LB plates containing the appropriate antibiotic selection for the plasmid.

3.10.3 Selection of Target Recombinants

3.10.3.1 Blue / White Colony Selection

This selection method was mainly used for cloning of target DNA in pGEM®-T Easy Vector. Blue/white colony selection is based on the ability of ß-galactosidase to produce a blue cleavage product from the artificial chromogenic substrat 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal) added to LB-plates.

3.10.3.2 PCR Colony Screening

This method bypasses DNA purification, and relies on the selectivity of PCR amplification to determine whether a bacterial colony of interest does indeed contain the desired DNA. Simply adding a small portion of a bacterial colony to a PCR master mix will introduce enough templates DNA for amplification. A single bacterial colony was added to a PCR tube containing the appropriate amount of water for a 25 µl amplification reaction using a sterile toothpick. The toothpick was agitated in the water to remove the colony. Subsequently, the suspension was put into a 5 ml LB-glass culture tube containing the appropriate antibiotic selection and centrifuged for 2-3 min to pellet the cell debris. An aliquot (5-10 µl) was used as the template in the PCR reaction. The remaining components were added to the PCR reaction

and subjected to normal cycling parameters for the particular primers. If insert orientation, as well as presence, needs to be determined, utilization of a forward vector-specific primer and a reverse insert-specific primer, or vice versa, allows such determination. If only the presence of the insert needs to be determined, then two insert-specific primers can be used. An additional 5 min denaturation step at 95°C before the amplification cycles will aid in lysing the bacteria to enhance PCR product amplification success. The resulting PCR products were checked on an agarose gel for the presence of the predicted band.

PCR reaction:

10x DNA polymerase buffer		2.5 µl
10 mM dNTPs		1 µl
Primers (forward and reverse)		2 µl
Polymerase		0.25 µl
H2O	up to	25 µl

PCR program:

95° C (denature)	5 min
50° C (anneal)	2 min
72° C (extend)	2 min
Repeat for	35 cycles

3.10.3.3 Restriction Analysis

In restriction analysis, recombinant DNA was treated with restriction enzymes to verify that a piece of DNA had been successfully inserted. A restriction map gives an indication of the relative locations of restriction sites in a DNA molecule. Recombinant colonies may be screened for the presence of the desired insert by restriction digestion of plasmid DNA from a mini preparation. In general, plasmid DNA from a 5 ml culture can be isolated and subjected to diagnostic restriction digestions. Restriction enzymes are also called restriction endonucleases. A DNA restriction endonuclease is an enzyme that recognizes a particular sequence of DNA bases and catalyzes the cleavage of the double-stranded DNA. The cuts in each strand may be opposite to one another, in which case the ends are said to be "blunt." More often, the cut in one strand was slightly displaced from that in the other one, leaving

"sticky" ends which can experience complementary base pair recognition binding. A restriction site is a site where cleavage by a DNA restriction endonuclease takes place. Vectors (such as plasmids) may be cut by restriction enzymes. The site can be reconnected or a piece of new DNA may be inserted.
The restriction analysis was performed in a 10 µl volume as follows:

Plasmid DNA		4 µl
10x Restriction buffer		1 µl
Restriction enzyme (1.7)		1 µl
H_2O	up to	10 µl

The reaction was incubated at 37°C for 1 h.

3.11 Isolation of Plasmid DNA

3.11.1 Isolation of Plasmid DNA (Mini-prep) by Alkaline lysis

This procedure which isolates rapidly plasmid DNA from bacteria uses the alkaline lysis method followed by isopropanol precipitation. Alkaline lysis is the method of choice for isolating circular plasmid DNA from bacterial cells, developed by Birnboin and Doly (1979) and Ish-Horowicz and Bruke (1981). Briefly, each single colony is inoculated into 5 ml of LB medium (50 µg/ml ampicillin) in glass culture tubes and grown overnight at 37°C with vigorous shaking (250 rpm). The next day, 1.5 ml culture was transferred into an Eppendorf tube and bacterial cells were harvested by centrifugation at 5000 rpm for 5 min. The pellet was resuspended in 250 µl of ice-cold Buffer P1 containing RNase (1 mg/10 ml). 250 µl of Buffer P2 were added, gently mixed, and the mixture was incubated at room temperature for 2-3 min. 250µl of ice-cold Buffer P3 was added and gently mixed. The mixture was incubated on ice for 5-10 min and then centrifuged at 14000 rpm for 10 min. The supernatant (800 µl) was transferred into a new Eppendorf tube and extracted twice with chloroform. Then 0.7 volumes of isopropanol were added to the supernatant. The suspension was vortexed and recentrifuged at 14000 rpm 15 min. The pellet was washed with 500 µl of 70% ethanol,

vortexed, and centrifuged for 10 min at 14000 rpm. The sample was dried for 30 min, and redissolved in 50 µl of sterile water or TE buffer. The DNA was stored at -20°C.

3.11.2 Plasmid Spin Mini-preparation

Plasmid isolation was performed using a spin mini-preparation with ready-to-use kits. Different kits were used. Plasmid purification protocols were based on a modified alkaline lysis procedure, followed by binding of plasmid DNA to an anion-exchange resin under appropriate low-salt and pH conditions. RNA, proteins, dyes, and low-molecular-weight impurities were removed by a medium-salt wash. Plasmid DNA was eluted in a high-salt buffer and then concentrated and desalted by isopropanol precipitation.

3.12 Gel Electrophoresis

3.12.1 Agarose Gel Electrophoresis

Electrophoresis is the migration of charged molecules in solution in response to an electric field. The rate of migration depends on the strength of the field, the net charge, size and shape of the molecules as well as the ionic strength, viscosity and temperature of the medium in which the molecules are moving. As an analytical tool, electrophoresis is simple, rapid and highly sensitive. It is used to study the properties of a single charged species and as a separation technique. In the case of DNA we can separate the molecules based on their size. DNA has a negative charge in solution, so it will migrate to the positive pole in an electric field. In agarose gel electrophoresis, the DNA is forced to move through a sieve of molecular proportions that is made of agarose, with large fragments of DNA moving more slowly than small fragments of DNA. The gel is prepared by dissolving agarose powder in electrophoresis buffer (TAE Buffer) at the desired concentration and heating in a microwave oven until complete melting. Ethidium bromide at a final concentration of 0.5 µg/ml is added to the gel at this point to facilitate visualization of DNA after electrophoresis. The solution was cooled to about 50°C. According to the number of samples to be analysed, the gel was poured into a 6 × 7 cm gel tray (Mini Sub Cell GT, BioRad) and allowed to solidify at room temperature. Electrophoresis was performed on an electrophoresis apparatus (BioRad) in 1x TAE buffer at

different voltage values (100-140 mV). Samples containing DNA were mixed with 6x loading buffer and pipeted into the sample wells. The lid and power leads were placed on the apparatus and a current was applied. To estimate the size of nucleic acids, a 100 bp DNA ladder was run in parallel on the same gel as the samples. DNA was migrated towards the anode which is usually colored red. To visualize DNA, the gel was placed on an UV-equipped transilluminator (Multiimage Light Cabinet, Biozyme).

3.12.2 Isolation of DNA Fragments from Agarose Gels

Isolation of a DNA fragment from an agarose gel was performed by using the NucleoSpin® Extract method, for which different kits were employed. DNA binds in the presence of chaotropic salts (e.g. the buffers NT1 and NT2) to a silica membrane. The buffer NT1 contains additional components in order to dissolve agarose gel slices. Afterwards, binding mixtures were loaded directly onto NucleoSpin® Extract columns. Contaminations like salts and soluble macromolecular components were removed by a simple washing step using ethanolic buffer NT3. Pure DNA was finally eluted under low ionic strength conditions with the slightly alkaline buffer NE (5 mM Tris-HCl, pH 8.5).

3.12.3 Sequencing Gel

The sequencing reaction yields strands that all have different lengths and end on a fluorescently labelled ddNTP. This mixture has to be separated in an acrylamide gel . While migrating through the gel the fluorescently labelled fragments pass a laser beam at the bottom of the gel. The laser excites the fluorescent molecule which sends out light of a distinct wavelength. This light is collected and focused by lenses into a spectrograph. For sequence gel composition, see (2.6.4).

3.12.4 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Gel electrophoresis is a useful method to separate and/or identify proteins and nucleic acids (Laemmli, 1970). In SDS-polyacrylamide gel electrophoresis (SDS-PAGE), proteins are separated largely on the basis of their molecular masses which can be estimated by this way.

SDS, however, denatures proteins and activity stains cannot be used to identify particular enzymes. Commonly, two sequential gels are used (2.5.4). The top gel, called the stacking gel, is slightly acidic (pH 6.8) and has a low (5.5%) acrylamide concentration to make a porous gel. Under these conditions, proteins separate poorly but form thin and sharply defined bands. The lower gel, called the separating or resolving gel, is more basic (pH 8.8) and has a higher polyacrylamide content (10 to 12%), which causes the gel to have narrower channels or pores. As a protein, concentrated into sharp bands by the stacking gel, travels through the separating gel, the narrower pores have a sieving effect, allowing smaller proteins to travel more easily and hence more rapidly than larger proteins.

3.12.4.1 Preparation of Gel and Samples

Polyacrylamide gels were formed by polymerization of the acrylamide monomer through cross-linking with N´-methylene bisacrylamide promoted by the addition of ammonium persulfate (APS) and TEMED (*N,N,N´,N´*-tetramethyl ethylenediamine). TEMED also acts as a stabilizing agent. Firstly, the applied protein mixture was concentrated in a highly porous stacking gel and then separated in the resolving gel. The composition of the gel mixture was used for preparing a 10 and 12% resolving gel. Immediately after addition of APS and TEMED, the mixture for the resolving gel was poured between two vertically oriented glass plates in a gel-casting stand to make a gel of 0.75 mm thickness. Protein samples to be analysed were mixed with protein loading buffer (ratio 1:2) and denatured at 95°C for 5 min, followed by loading the samples into the wells of the stacking gel.

3.12.4.2 Electrophoresis and Detection of Protein Bands

Electrophoresis was carried out on a Minigel-Twin Electrophoresis System (Biometra) in 1x SDS electrode buffer at 200 V, 75 mA and 15 W supplied by a Multidrive XL-power supply (Pharmacia) for approximately 90 min.
The protein bands were visualized by two methods:

1- Coomassie Blue staining:

Incubate the gel in Coomassie Blue staining solution for 60 min or overnight, followed by destaining for 60 min.

2- Silver staining:

Process the gel according to the procedure given in Table 3-1, with 250 ml of each solution being used per gel. The timing of the individual steps differs according to the type of solution. Shake slowly on a shaker or rocker.

Tab. 3-2: Silver staining protocol

Solution	Time for gel incubation
Fixing solution	40-60 min
Sensitizing solution	60 min or overnight
Distilled water	3 x 5 min
Silver solution	30 min
Distilled water	2 x 2 min
Developing solution	Until protein bands are visible (5-10 min)
Stopping solution	10 min

3.12.4.3 Determination of Molecular Masses

Generally, SDS-PAGE is used for proteins and PAGE or agarose gel electrophoresis for nucleic acids. Samples of known molecular masses are run along with the protein or nucleic acid to be characterised. A linear relationship exists between the logarithm of the molecular mass of an SDS-denatured polypeptide, or native nucleic acid, and its R_f. The R_f is calculated as the ratio of the distance migrated by the molecule to that migrated by a marker dye-front. A

simple way of determining a relative molecular mass (M_r) by electrophoresis is to plot, for known samples, a standard curve of distances migrated vs. the log10 molecular masses and to read off the logM_r of the unknown sample after measuring the distance migrated on the same gel.

3.13 Sequence Analysis

Plasmid DNA from selected clones was isolated for sequencing following a mini-alkaline lysis precipitation procedure. Both strands of the insert were sequenced using the T7 and SP6 primers or synthetic oligonucleotide primers, as needed to extend the sequence. Sequencing reactions were carried out in an ABI PRISM® 377 DNA Sequencer (Applied Biosystems). This DNA Sequencer automatically analyzes DNA molecules labeled with multiple fluorescent dyes. After loading samples onto the system's vertical gel, they undergo electrophoresis, laser detection, and computer analysis. All sequences were edited and analyzed with Lasergene 99 DNASTAR Software. Database searches for sequence homology and comparisons were performed with BLAST analytical tools. Recently, sequence analysis involved the company MWG Biotech.

3.13.1 Sequencing Reaction Protocol

The sequencing reactions were carried out with Taq-DNA-Polymerase and the ABI PRISM®–BigDye Terminator Cycle Sequencing Ready Reaction Kit from Applied Biosystems (Weiterstadt) according to the manufacturer's protocol. This kit contains the four ddNTPs with different fluorescence labels (BigDye Terminators).

The sequencing reaction was as follows:

BigDye solution		2 µl
DNA template		500-700 ng
Primer		1 µl
Distilled water	up to	10 µl

PCR tubes were subjected to 30 cycles in a thermocycler using the following program:

1. Denaturation	15 s at 95°C	
2. Annealing	25 s at 50°C	30 cycles
3. Elongation	1 min 30 s at 72°C	

These PCR steps have the following function:

1- The template DNA is denatured.

2- A unique primer anneals and second strand synthesis begins. Dye-labeled chain-terminating nucleotides (ddNTPs) are randomly incorporated.

3- Fragments of different lengths are created. Each dye-labeled nucleotide that was incorporated terminates the strand synthesis.

4- The DNA fragments were separated according to size on a polymer-based matrix in the presence of an electric current. Based on the fluorescent tag, the instrument will assign the base calls. After sequencing, the reactions must be cleaned up to remove the unincorporated fluorescent nucleotides.

3.13.2 Sequencing Reaction Precipitation

Cleanup of sequencing reaction samples by ethanol precipitation was carried out as following: 1.5 ml tubes were prepared by adding sodium acetate/EDTA buffer and 90 µl water. The full sequencing reaction was added to the tube and all components were mixed. After addition of 100% ethanol and vortexing, the tubes were centrifuged for 15 min at ~15000 rpm at room temperature, and the supernatant was discarded. The tubes were rinsed with 250 µl of 70% ethanol and centrifuged. The supernatant was poured off and the rest of ethanol was removed by centrifugation. The pellet was dried and stored at -20C.

3.13.3 Computer-assisted Sequence Analysis

The following programs were used to analyse and evaluate newly obtained sequence information:

BLAST (**B**asic **L**ocal **A**lignment **S**earch **T**ool):

The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. BLAST can be used to infer functional and evolutionary relationships between sequences as well as to help identify members of gene families. This program is accessible at: **http://www.ncbi.nlm.nih.gov/BLAST/**

DNAStar Lasergene

This is a package of computer programs for manipulating and analyzing DNA and protein sequence data and includes:

SeqMan: To assemble fragment data from sequencing projects of any size - from a few fragments to tens of thousands of fragments

GeneQuest: To locate genes, regulatory elements, patterns and structures in your primary DNA sequence.

Protean: To predict and display patterns, secondary structural characteristics and physicochemical properties of protein sequences via its comprehensive suite of protein analysis tools.

MegAlign: Offers a choice of four pair-wise and four multiple sequence alignment methods for aligning nucleic acid or polypeptide sequences.

PrimerSelect's: Comprehensive set of tools enabling you to design and analyze primers for PCR, sequencing, probe hybridization and transcription.

EditSeq: To work on nucleic acid and protein sequences of all sizes from a wide variety of popular formats. You can also utilize the integrated internet interface to search NCBI's databases to locate sequences by accession number or sequence similarity via BLAST.

3.14 Heterologous Expression and Protein Purification

The purification of a recombinant protein is greatly accelerated if the protein can be isolated from cells overproducing it. To maximize expression, the cloned gene must be transcribed and translated as efficiently as possible. This is feasible due to the construction of expression vectors, i.e. modified plasmids with useful features which can be propagated and controlled in special hosts (expression systems). Usually, vectors for cloning and expressing target DNA are derived from medium-copy plasmids such as pRSET-B. *E.coli* expression systems should meet several criteria including (i) minimal basal expression of the gene to be expressed under repressed conditions, (ii) fast and uncomplicated induction of a wide variety of genes to a

high level of expression, and (iii) easy cloning and DNA manipulation features. The most common expression system is the T7 expression system derived from bacteriophage T7. The T7 expression system is based on the use of the T7 bacteriophage promoter and RNA polymerase. The T7 RNA polymerase is useful for synthesizing selectively large amounts of RNA because the T7 RNA polymerase recognizes only the T7 promoter and not the *E.coli* promoters. Conversely, the *E.coli* RNA polymerase does not recognize the T7 promoter. The T7 RNA polymerase is able to transcribe genes five times faster than the *E.coli* RNA polymerase (Chamberlin and Ring, 1973; Golomb and Chamberlin, 1974). The gene encoding the T7 RNA polymerase was inserted into the chromosome of the bacteria used for overexpression. The expression of this gene is under the control of the lacUV5 promoter. Expression of the target gene was induced by addition of IPTG (isopropyl-ß-D-thiogalactopyranoside) to a growing culture.

3.14.1 Construction of pRSET Expression Vector

PCR was used to introduce NheI and EcoRI sites at the 5`and 3`ends, respectively, of a cDNA. The amplified product was digested with NheI and EcoRI and cloned into the same sites of pRSET-B vector containing a histidine-tag. The expression vector was transferred into an *E.coli* strain and the growth of bacterial cells and the induction of expression with IPTG were preformed.

Digestion of DNA insert and vector pRSET B:

1 µg purified DNA insert or pRSET-B vector		3 µl
10x Tango Buffer		3 µl
10 U/µl EcoRI		1 µl
20 U/µl NheI		0.5 µl
H_2O	up to	25 µl

The reaction was incubated at 37°C for 2 h.

3.14.2 Confirmation of Insert

After isolation, the plasmids were analyzed for the presence of an insert into the multiple cloning site (MCS). This was done by sequence analysis of the insert across the MCS using pRSET forward and reverse primers (Invitrogen) which flank the multiple cloning site of pRSET-B.

3.14.3 Growing and IPTG treatment of *E.coli* BL21

A single isolated bacterial colony from freshly streaked plates (grown on LB agar medium containing 50 mg /ml ampicillin and 34 mg/ml chloramphenicol) was used to inoculate 10-ml liquid cultures in LB medium containing the same concentrations of the two antibiotics. Growth was overnight with shaking at 225 rpm at 37°C. One 5 ml-aliquot of each culture was used to inoculate 50-ml liquid cultures containing 50mg/ml ampicillin. Once the cultures reached OD_{600} 0.4-0.5, recombinant protein expression was induced by the addition of isopropyl -ß-D-thiogalactopyranoside (IPTG), and the culture was grown for 20–24 h at 25°C with shaking at 150 rpm. Cells were harvested by centrifugation at 6000 rpm for 10 min at 4°C. Pellets were resuspended in 10 times the volume with 50 mM sodium dihydrogen phosphate buffer at pH 7.0. Cells were disrupted by sonication on a Sonifer Cell Disruptor on ice. The resulting homogenate was centrifuged at 15000 rpm for 20 min to pellet the debris. The supernatant was assayed for activity and stored at -20°C prior to protein purification.

3.14.4 Preservation of bacteria

Bacterial cultures containing plasmids with target insert were stored in glycerol. In an Eppendorf tube, 800 µl of bacterial culture were mixed with 200 µl of glycerol (sterilized by autoclaving). The mixture was vortexed to ensure that the glycerol is dispersed. The tube was stored at -80°C for long-term storage. To recover the bacteria, the frozen surface of the culture was scraped with a sterile inoculating needle and the bacteria adhered to the needle were immediately streaked on the surface of a LB agar plate containing the appropriate antibiotic. The plates were incubated overnight at 37°C.

3.14.5 Expression and Purification of Recombinant Proteins

Many natural proteins have metal binding sites which can be used for purification. The concept of this type of purification tool is rather simple. A gel bead is covalently modified to display a chelator group for binding a heavy metal ion like Ni^{2+}. Affinity chromatography is viewed as a group-selective tool for purifying the metal-binding class of proteins. His-tagged recombinant protein can be purified by Metal Chelate Affinity Chromatography. The initial stage of His-tagged protein purification is based on the remarkable selectivity and high affinity of patented Ni-NTA (nickel-nitrilotriacetic acid) resin for proteins containing an affinity tag of six consecutive histidine residues, the 6xHis-tag. NTA, which has four chelating sites for nickel ions, binds nickel more tightly than metal-chelating purification systems that have only three sites available for interaction with metal ions. The extra chelation site prevents nickel ion leaching, providing a greater binding capacity and high-purity protein preparations. Isolation and purification of proteins were conducted using His-trap columns (Amersham Biosciences). Fusion proteins were purified directly from bacterial lysates and recovered from the matrix under mild elution conditions using imidazole, which preserves antigenicity and functionality of the protein.

3.14.6 Fast Protein Liquid Chromatography (FPLC) Analysis

The FPLC system is a liquid chromatography for rapid purification of proteins and other biomolecules. Biomolecules have various characteristics such as molecular mass, electric charge, and hydrophobicity. As so, purification of the object is usually achieved by using a combination of chromatographic methods including gel filtration, ion exchange, hydrophobic interaction, and affinity chromatography.

3.14.6.1 FPLC Procedure

The crude protein was loaded onto a 5 ml-nickelnitrilotriacetic acid column that had been pre-equilibrated with washing buffer (50 mM sodium phosphate, pH 7.0, 0.5 M NaCl, 20 mM imidazole). After loading, the column was washed with three column volumes of washing buffer and bound protein was eluted by application of three column volumes of washing buffer supplemented with a further 1 M imidazole (elution buffer).

FPLC Program

Tab. 3-3: FPLC Program

Step Number	(ml)	Process
1	25	Collect fraction within 1 time window ending at 40.0 ml
2	0	Turn UV lamp ON
3	0	Isocratic flow with 100% washing-buffer 1, 0% elution buffer at 1 ml/min
4	5	Static loop: inject 5.0 ml sample at 0.50 ml/min for 10 min
5	10	Isocratic flow with 100% washing buffer, 0% elution buffer at 1 ml/min for 15 ml
6	25	Linear gradient with 0% to 100% elution buffer at 1ml/min for 10 ml
7	35	Isocratic flow with 0% washing buffer, 100% elution buffer at 1ml/min for 5 min
8	40	Isocratic flow with 100% washing buffer, 0% elution buffer at 1ml/min for 5 ml
9	45	Turn UV lamp OFF
10	45	End of protocol

3.14.6.2 Buffer Change and Desalting of Protein Samples

Eluted protein was applied immediately onto a 10-ml PD_{10} desalting column (Amersham Biosciences) that had been previously equilibrated with five column volume of 0.16 M NaOH with desalting buffer and then five column volumes of distilled water. A PD_{10} column was used to remove the imidazole after the final purification step.

3.14.6.3 Protein Determination

Protein concentrations were measured by the method of Bradford (1976) using bovine serum albumin (BSA) as standard. To obtain a calibration curve, aliquots containing 1 - 10 µg BSA were mixed with 900 µl freshly prepared Bradford dye solution. The volume was made up to 1 ml with water and the contents in each cuvett were mixed carefully. After 5 min, the absorbance of the samples at 595 nm was measured using a blank of 900 µl Bradford solution with 100 µl water. The same procedure was applied to aliquots of purified protein, and protein concentrations were determined from the calibration curve.

3.14.7 Gel filtration

For molecular mass estimation, a HiPrep16/60 Sephacryl S-200 High Resolution column was used. Determination of the molecular mass of a new protein requires preceding calibration of the column with standard proteins. The following markers were used:

- Albumin, Bovine serum, 66 kDa
- Alcohol dehydrogenase, Yeast, 150 kDa
- β-Amylase, Sweet Potato, 200 kDa
- Blue dextran, 2000 kDa
- Carbonic anhydrase, Bovine erythrocytes, 29 kDa
- Cytochrome c, Horse heart, 12.4 kDa

The column was equilibrated with 50 mM Tris –HCl, pH 7.5 containing 100 mM KCl.
To determine the void volume (V_o), blue dextran (2000 kDa) was dissolved in equilibration buffer (2 mg/ml) and passed through the gel filtration column at a flow rate of 0.3 ml/min. The fraction volume was 5 ml. To determine the elution volume (V_e) of the standards, the proteins were dissolved in equilibration buffer at the following concentrations (Tab. 3-4):

Recommended concentrations of standard proteins:

Tab. 3-4: Recommended concentrations of standard proteins

Standard protein	Concentration used
Albumin	10 mg/ml
Alcohol dehydrogenase	5 mg/ml
β-Amylase	4 mg/ml
Carbonic anhydrase	3 mg/ml
Cytochrome c	2 mg/ml

The individual samples (2 ml) were applied to the column. A standard curve was calculated by plotting the molecular mass vs. V_e/V_o for each standard protein. The relative molecular mass of a new protein was concluded from the standard curve.

3.15 Enzyme Assay

3.15.1 Spectrophotometric Assay

Enzyme activity was assayed by monitoring the formation of CoA esters of various cinnamate derivatives, as described by Knobloch and Hahlbrock (1977). The reaction mixture contained the enzyme solution, 5 mM ATP, 5 mM $MgCl_2$, 0.2 mM CoA, and 0.2 mM cinnamate derivative in a total volume of 1 ml. Two control mixtures lacking either CoA or enzyme solution was used. Enzyme activity was measured by following the increase in absorbance at the absorption maximum of the appropriate CoA ester (Lee et al., 1997).

3.15.2 High performance liquid chromatography (HPLC) Assay

The assay was carried out as described by Barillas and Beerhues (1997) in 1.5 ml Eppendorf tubes at 35°C for 1h in a water bath. After incubation, the reaction was stopped by addition of 10 µl 3 M trichloroacetic acid which denatured and precipitated the protein. Incubation reactions were centrifuged at 13.000 rpm for 10 min, the supernatant was subjected to HPLC analysis.

3.15.2.1 Instrumental

HPLC Breeze
Pumps: Waters 1525 Binary HPLC Pump
Detector: Waters 2487 Dual Absorbance
Software: Waters Breeze GPC
Column: C_{18} 100-5 (25 x 0.4 cm; Macherey-Nagel, Düren, Germany)

HPLC (photo diode array detection)

Injector: Waters712 auto sampler
Detector: Waters 991 photo diode array
Pumps: Waters
Controller: Waters 600E system
Column: C_{18} 100-5 (25 x 0.4 cm; Macherey-Nagel, Düren, Germany.

3.15.2.2 HPLC gradients

1. Methanol - Water System

Tab. 3-5: HPLC gradients (Methanol - Water System)

Gradient	Time [min]	Methanol [%]	Water [%]	Wave length [nm]
1 Flow rate 0.5 ml/min	0 3 6 8 16 22 24 27 28 30	30 30 40 45 50 60 100 100 30 30	70 70 60 55 50 40 0.0 0.0 70 70	261-333nm
2 Flow rate 1 ml /min	0 3 8 18 33 40 42 45 47 50	10 10 20 20 50 60 100 100 10 10	90 90 80 80 50 40 0 0 90 90	261-333 nm

2. Acetonitrile - Water System

Tab. 3-6: HPLC gradients (Acetonitrile - Water System)

Gradient	Time [min]	Acetonitrile [%]	Water [%]	Wave length [nm]
1 Flow rate 0.5 ml/min	0 5 22 25 28 30 32 35	5 5 38 75 100 100 5 5	95 95 62 25 0 0 95 95	261-333 nm

3. Acetonitrile – Potassium phosphate system

Tab. 3-7: HPLC gradients (Acetonitrile – Potassium phosphate system)

Gradient	Time [min]	Acetonitrile[%]	Water + 3% potassium phosphate [%]	Wave length [nm]
1	0	50	50	261-333 nm
	2	50	50	
Flow rate 0.5	32	80	20	
ml/min	37	100	0	
	40	100	0	
	42	50	50	
	47	50	50	

4. Isocratic elution system

The composition of the solvent remains constant throughout the analysis.

4. RESULTS

4.1 Degenerate Oligonucleotide Primer PCR (DOP-PCR)

As a working hypothesis, benzoate:CoA ligase and cinnamate:CoA ligase were assumed to be related to 4-coumarate:CoA ligase (4CL). The following 4CL sequences publically accessible from the National Center for Biotechnology Information were aligned to generate a 4CL amino acid consensus sequence (Fig. 4-1).

The alignment of these amino acid sequences indicated that, although the N-terminal region of 4CLs is highly variable, large stretches of the proteins are highly conserved and thus could be used as primer targets. Three consensus regions were used to derive one degenerate sense and two degenerate antisense primers (Fig. 4-1). The sense primer was located at the first putative AMP-binding domain (**LPFSSGTTGLP**) and the two antisense primers were designed based on the second putative AMP-binding domain (**GEICIRG**) and the motif (**GQGYGMTEA**). The nucleotide sequences of the degenerate primers are given under 2.8.3.

4.2 Amplification of *Hypericum androsaemum* cDNA Fragments

A cDNA library complementary to mRNA from 3-day-old *H. androsaemum* cell cultures (Schmidt and Beerhues, 1997) was available in the working group. Recombinant λDNA was prepared from a phage lysate (3.1.3) and used for PCR, together with combinations of forward and reverse primers or of one degenerate primer and the M13 universal primer (Fig. 4-2). The amplified fragments ranged in size from 450 to 470 bp. Their nucleotide sequences exhibited 37 to 57% identity to the corresponding regions of known 4CLs.

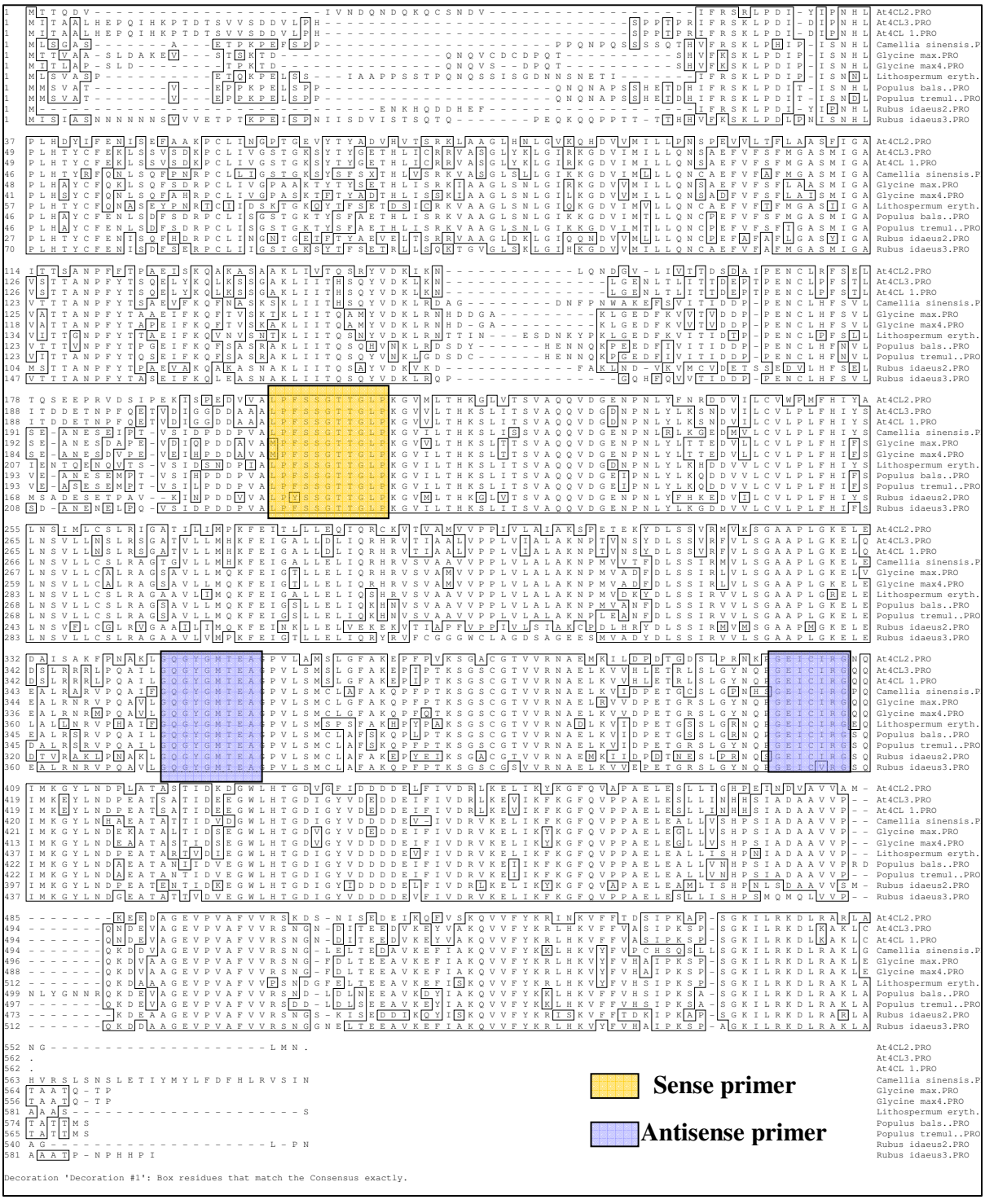


Fig. 4-1: Alignment of deduced 4CL amino acid sequences. Boxes indicate identical amino acids in three or more sequences and colored boxes indicate the locations of the degenerate primers.

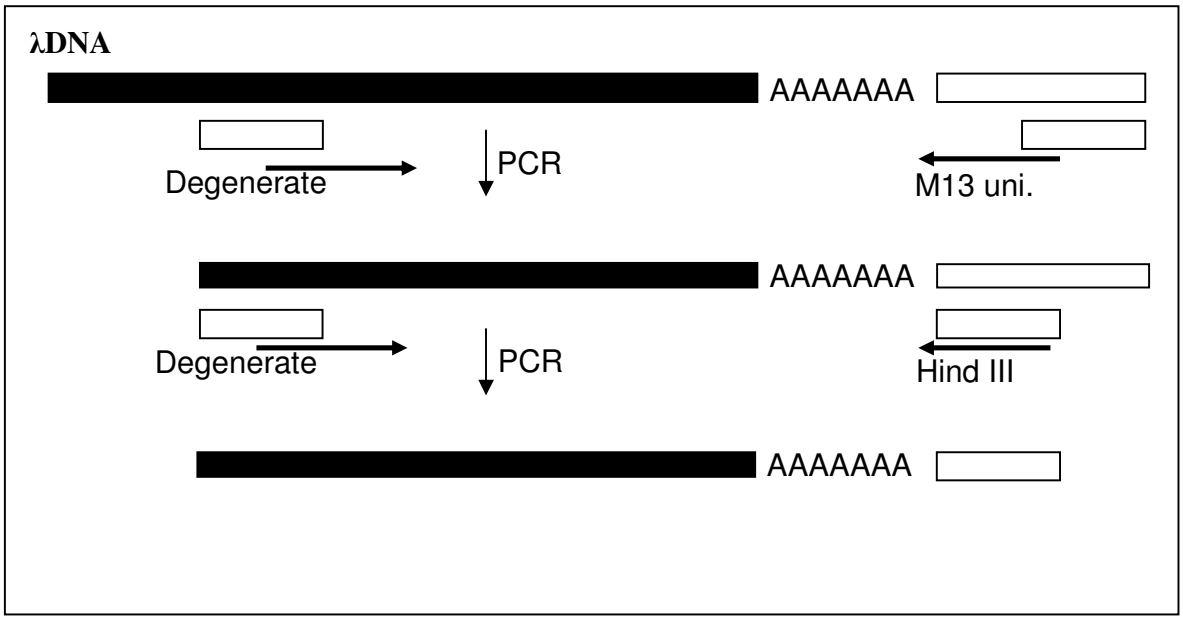


Fig. 4-2: Scheme of PCR with degenerate and lambda phage specific primers

The following λDNA fragments were obtained:

- **Fragment 1:**

This λDNA fragment was amplified using the primers Forward 1 and CoA ligase reverse.

- **Nucleotide sequence of fragment 1:**

CTCTATTCCTCGGGGACGACTGCTGGGAGTTAACGTGAGAACTTGACACACGAGATCTTGAATTCTATGGATGAACTT
ATGCTGAGGGTTGAGGCGTCTCTGTACAAGGATTGGCCTAGTTCAGAGAATTCGTATAATGCTACTGTCCCGATGTTT
CACATATACGGTATAGCGTATATCGTCATGGCGCTGTTGTCACTGGCGACTTGTGATGATGTGATGAAGACGTTCAAT
CCGGATGAGGCCGTGTAGGTTATCTGAGATAGGTACCGCGTGACTCATCTTTCGGTGGATCCACCGATGTTGTCTTCG
ATGATCAGGAGTGCGAAAAATTGTGGAAGTCCCTTGAAGAGCTTGATACAGAATTGTTCCGCTGCAGCCGGTCTCTCT
GGCAAGGCGATCGAGGAATTCTAGGAAGCTTATCCTCATGTTGATATCATTCAAGAATACGGCATGACCGAAGCG

- **Amino acid sequence and data bank comparison of fragment 1:**

```
gi|15234087|ref|NP_193636.1| 4-coumarate-CoA ligase [Arabidopsis thaliana]
 Length=566

             Identities = (52%)     Frame = +1

Query  1    LYSSGTTAGS*RENLTHEILNSMDELMLRVEASLYKDWPSSENSYNATVPMFHIYGIAYI  180
            +YSSGTT  S    LTH  L +  EL +R EAS Y ++P S N Y A +P+ HIYG++
Sbjct  210  MYSSGTTGASKGVLLTHRNLIASMELFVRFEASQY-EYPGSSNVYLAALPLCHIYGLSLF  268

Query  181  VMALLSLATCDDVMKTFNPDEAV*VI*DRYRVTHLSVDPPMLSSMIRSAKN-CGSPLKSL  357
            VM LLSL +   VMK F+  + V VI +R+++TH  V PPML ++ + AK  CG   KSL
Sbjct  269  VMGLLSLGSTIVVMKRFDASDVVNVI-ERFKITHFPVVPPMLMALTKKAKGVCGEVFKSL  327

Query  358  IQNCSAAAGLSGKAIEEF*EAYPHVDIIQEYGMTEA  465
             Q  S AA LS K IE+F +  PHVD+IQ YGMTE+
Sbjct  328  KQVSSGAAPLSRKFIEDFLQTLPHVDLIQGYGMTES  363
```

- **Fragment 2:**

This fragment resulted from amplification with Forward 1 and CoA ligase reverse.

- **Nucleotide sequence of fragment 2:**

```
CTCTACTCTTCGGGTACTACTGGTGCGAGTAAAGGTGTGATCTTGGCACACGAGAACTTGATTTCTATGGTTGAACTT
ATGGTGAGGGTTGAGGCCTCTCTGTACAAGGATTGGCCTAGTTCAGAGAATTCGTTTATAGCTACTGTCCCGATGTTT
CCACATAcTACGGTATAGCGTATTTCGTCATGGGGCTGTTGTCACTGGGGACYTGTGTTGTTGTGATGAAGAGGTTCA
ATCCGGATGAGGCCGTGACGGTTATCGATAGGYACCGCGTGACTCATCYTACGGTGGATCCCACCGAAGTTGTCAGCG
ATGATCAGGAGTGCGAAAAATTGTGGAARTCCCATGAAGAGCTTGAAACAGATTTGTTGCGGTGCAGCCCCTGTCTCT
GCCAAGGCCATCCAGGAATTCTTGGAAGCAATTCTTCATGGTGATACAAAACACGGGCACTGCATGACAGAGTCT
```

- **Amino acid sequence and data bank comparison of fragment 2:**

```
gi|15234087|ref|NP_193636.1| 4-coumarate-CoA ligase [Arabidopsis thaliana]
Length=566
             Identities = (37%)  Frame = +1

Query  1    LYSSGTTGASKGVILAHENLISMVELMVRVEASLYKDWPSSENSFIATVPMFPHTTV*RI  180
            +YSSGTTGASKGV+L H NLI+ +EL VR EAS Y ++P S N ++A +P+
Sbjct  210  MYSSGTTGASKGVLLTHRNLIASMELFVRFEASQY-EYPGSSNVYLAALPL---------  259

Query  181  SSWGCCHWGXVLL**RGSIRMRP*RLSIGTA*LILR--------------------WIPP  300
                 CH    + L    G         LS+G+  ++++                    +PP
Sbjct  260  -----CHIYGLSLFVMG-------LLSLGSTIVVMKRFDASDVVNVIERFKITHFPVVPP  307

Query  301  KLSAMIRSAKN-CGXPMKSLKQICCGAAPVSAKAIQEFLEAILHGDTKHGHCMTES  465
             L A+ + AK  CG   KSLKQ+  GAAP+S K I++FL+ + H D   G+ MTES
Sbjct  308  MLMALTKKAKGVCGEVFKSLKQVSSGAAPLSRKFIEDFLQTLPHVDLIQGYGMTES  363
```

- **Fragment 3:**

This fragment resulted from amplification with Forward 1 and CoA ligase reverse.

- **Nucleotide sequence of fragment 3:**

GGTGCGAGTAAAGGTGTGATCTTGACACACGAGAACTTGATTTCTATGGTTGCACTTATGGTGAGGGTTGAGGCCTCT
CTGTACAAGGATTGGCCTAGTTCAGAGAATTCGTTTATAGCTACTGTCCCGATGTTTCACATTATACGGTATAGCGTA
TTTCGTCATGGGGCTGTTGTCACTGGGGACTTGTGTTGTTGTGATGAAGAGGTTCAATCCGGATGAGGCCGTGAAGGT
TATCGATAGGTACCGCGTGACTCCATCTTTCGGTGGTTCCACCTGATGTTGTCTTCGATGATCAGGAGTGCGAAAAAT
TGTGGAAGTCCCTTGAAGAGCTTGAAACAGATTTGTTGCGGTGCAGCCCCTCTCTCTGCCAAGGCCATCCAGGAATTC
TTGGAAGCTTTTCCTCATGTTGATTTCATTCAAGGCTACGGCATGACCGAAGCC

- **Amino acid sequence and data bank comparison of fragment 3:**

gi|46805609|dbj|BAD17022.1| **putative 4-coumarate-CoA ligase 4CL2 [*Oryza sativa (japonica cultivar-group)*]** Length=591

Identities = (**41%**) Frame = +1

```
Query  1    GASKGVILTHENLISMVALMVRVEASLYKDWPSSENSFIATVPMFHIIRYSVFRHGAVVT  180
            G SKGV+LTH NLI+M  L VR EAS Y    + EN ++A +PM H+   S+F  G +
Sbjct  233  GRSKGVVLTHRNLIAMTELFVRFEASQYHARGARENVYMAALPMSHVYGLSLFAVGLLSI  292

Query  181  GDLCCCDEEVQSG*GREGYR*VPRDSIFRWFHL-----MLSSMIRSAKNCGSP---LKSL  336
            G          +G        V     ++  H+     ++++M+R+A   G P   + SL
Sbjct  293  GATVVVMRRFDAGDA------VAAIGRYKVTHMPLVPPIMAAMVRAAAAGGVPPSQVASL  346

Query  337  KQICCGAAPLSAKAIQEFLEAFPHVDFIQGYGMTEA  444
             Q+ CGAAP++A  I EFL+AFPHVDFIQGYGMTE+
Sbjct  347  VQVSCGAAPITAALIHEFLQAFPHVDFIQGYGMTES  382
```

Fig. 4-3: Alignment of the *H. androsaemum* cDNA core fragments 1, 2, and 3 amplified with the CoA ligase reverse and forward 1 primers

Tab. 4-1: Percent identity of the *H. androsaemum* cDNA core fragments 1, 2, and 3 amplified with the CoA ligase reverse and forward 1 primers

Fragment comparison	Percent identity
Fragment 1 with Fragment 2	80
Fragment 1 with Fragment 3	86
Fragment 2 with Fragment 3	88.5

4.3 Gene Specific Primers used in *Hypericum androsaemum*

In order to amplify full-length cDNAs, gene specific primers were designed based on the sequences of the three core fragments (Fig. 4-3) and combined with λDNA specific primers for PCR under various conditions. Unfortunately, amplifications resulted in fragments of only 100 to 470 nucleotides in length. The isolation of full-length cDNAs turned out to be difficult. Despite repeated attempts, no clones containing inserts larger than 470 nucleotides were generated. The following fragment resulted from amplification with a gene specific primer based on the above fragment 1 and M13 universal. Only few amino acids were extended, as highlighted at the 5` end.

- **Nucleotide sequence :**

```
AAATCCTATTAATTCCTTAAGATCCCTGTTGGTTTGTTTAGGAAGCCGTTGATTACACAGGAAGATTCTGCAGCTATA
CTCTACTCTTCTGGTACTACTGGTGCGAGTAAAGGTGTGATCTTGGCACACGAGAACTTGATTTCTATGGTTGAACTT
ATGGTGAGGGTTGGGGCCTGTCTGTACAAGGATTGGCCTAGTTCACAGAATTCGGTTATTGCTACTGTACCGATGTGG
AACCACCTACGGTATAGCGTATTTCGTCATGGG
```

- **Amino acid sequence and data bank comparison:**

```
gi|15234087|ref|NP_193636.1| 4-coumarate-CoA ligase [Arabidopsis thaliana]
Length=566

            Identities = (49%)    Frame = +1

Query  1    KSY*FLKIPVGLFRKPLITQEDSAAILYSSGTTGASKGVILAHENLISMVELMVRVGACL  180
            K Y  +K   G   KPLI Q+D AAI+YSSGTTGASKGV+L H NLI+ +EL VR  A
Sbjct  184  KFYSIMKESFGFVPKPLIKQDDVAAIMYSSGTTGASKGVLLTHRNLIASMELFVRFEASQ  243

Query  181  YKDWPSSQNSVIATVPMWNHLRYSVFRHG  267
            Y ++P S N  +A +P+ +    S+F  G
Sbjct  244  Y-EYPGSSNVYLAALPLCHIYGLSLFVMG  271
```

4.4 *Sorbus aucuparia*

Due to the difficulties encountered in *H. androsaemum*, another model plant was used to continue the cloning efforts. Similar combinations of degenerate primers as designed above were used to isolate core cDNA sequences from cell cultures of *Sorbus aucuparia* established in our working group.

4.4.1 Isolation of mRNA and Amplification of Core Fragments

Pools of total RNA and mRNA were isolated from *S. aucuparia* cell cultures (3.1.1; 3.1.2) and used to synthesize cDNA by RT-PCR using the SMART RACE cDNA Amplification Kit (3.2.1). The core fragments obtained ranged in size from 400 to 650 bp and exhibited 62 to 94 % nucleotide sequence identity with 4CL genes from other species. Following the RT-PCR step, cDNA was used immediately for subsequent PCR with combinations of degenerate primers.

4.4.1.1 Core Fragments Amplified with the CoA ligase Reverse and Forward 1 Primers

The following three fragments were obtained:

- **Nucleotide sequence of fragment 1:**

```
TTGCCGTATTCGTCGGGGTTTACGGGGCTACCTAAAGGGGTTATGCTGACGCACAAAGGGTTGGTGACGAGCGTGGCG
CAACAGGTGAACGGAGAGAATCCGAATTTGTATTTCCACAGTGAGGACGTGATCCTCTGCGTGCTGCCCTTGTTCCAT
ATCTACTCCCTCAATTCAGTGTTTCTNTGCGGACTCAGAGTTGGGGCGGCGATACTGATCATGCAGAAGTTTGAGATC
ACCAAGTTGTTGGAGCTGGTGGAGAATTACAAGGTGACGATTGCGCCTTTTGTACCTCCGATCGTTTTGAGTATTGCC
AAAAGCCCCGACTTAGATCGGTACGACTTGTCATCGATAAGGATGGTGATGTCCGGTGCGGCGCCGATGGGGAAGGAG
CTTGAGGATACAGTGAGGGCTAAGTTACCTAGTGCCAAACTTGGACAAGGCTACGGCATGACCGAGGCCGAATCACTA
GT
```

- **Amino acid sequence and data bank comparison of fragment 1:**

gi|9651915|gb|AAF91309.1| **4-coumarate:CoA ligase 2 [*Rubus idaeus*]**
Length=544

```
            Identities = (94%)          Frame = +1

Query  1    LPYSSGFTGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHSEDVILCVLPLFHIYSLNSVF  180
            LPYSSG TGLPKGVMLTHKGLVTSVAQQVDGENPNLYFH EDVILCVLPLFHIYSLNSVF
Sbjct  188  LPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHKEDVILCVLPLFHIYSLNSVF  247

Query  181  LCGLRVGAAILIMQKFEITKLLELVENYKVTIAPFVPPIVLSIAKSPDLDRYDLSSIRMV  360
            LCGLRVGAAILIMQKFEI KLLELVE  KVTIAPFVPPIVLSIAK PDL RYDLSSIRMV
Sbjct  248  LCGLRVGAAILIMQKFEINKLLELVEKEKVTIAPFVPPIVLSIAKCPDLHRYDLSSIRMV  307

Query  361  MSGAAPMGKELEDTVRAKLPSAKLGQGYGMTEA  459
            MSGAAPMGKELEDTVRAKLP+AKLGQGYGMTEA
Sbjct  308  MSGAAPMGKELEDTVRAKLPNAKLGQGYGMTEA  340
```

- **Nucleotide sequence of fragment 2:**

```
GTAGTGATTTTGCCGTACTCGTCSGGGACGACAGGCTTGCCTAAGGGGGTGATGCTAACGCACAAGGGTCTTGTGACA
AGCGTTGCTCAGCAGGTGGACGGGGAAAATCCCAACTTGTATTATAGCACCGACGACGTCGTCTTATGCGTGCTGCCA
CTTTTTCATATATATTCTTTGAACTCGGTATTGCTTTGTGGACTTAGAGCCGGAGCTGCCATTTTGATGATGAACAAG
TTTGAGATTGTTTCTCGTTAGGGGTaTGATCGACATTACAAGGTTAGTATTGCCCCGATCGgTGCCtGCCGATAGTGT
cTGGCCATTGMCAAGTTTCCCGATCTTGATAAGTACGATaTTGTCGTCAATTCGAGTGCTTcAWGTGTGGWGGGGCAC
CSCTTGGGAASGAGCTTGAKGATRCTCtGAAcAGCCAYGCTTCCCATTGTCACACTTGGTCAGGGATATGGTATGACA
AACGCAGG
```

- **Amino acid sequence and data bank comparison of fragment 2:**

gi|9651917|gb|AAF91310.1| **4-coumarate:CoA ligase 1 [*Rubus idaeus*]**
Length=543

```
             Identities = (70%)       Frame = +1

Query  1    VVILPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYYSTDDVVLCVLPLFHIYSLN  180
            VV LPYSSGTTGLPKGVMLTHKGLVTSV+QQVDGENPNLYYS+DDVVLCVLPLFHIYSLN
Sbjct  184  VVALPYSSGTTGLPKGVMLTHKGLVTSVSQQVDGENPNLYYSSDDVVLCVLPLFHIYSLN  243

Query  181  SVLLCGLRAGAAILMMNKFEIVSR*GYDRHYKVSIAPIGACR*CLAIXKFPDLDKYDI--  354
            SVLLCGLRAGAAIL+M KFEIVS     + ++VS+API      LAI KFPDLDKYD+
Sbjct  244  SVLLCGLRAGAAILLMQKFEIVSLLELMQKHRVSVAPIVPPT-VLAIAKFPDLDKYDLGS  302

Query  355  --VVNSSASXVXGHXLGXSLXXL*TAXLPIVTLGQGYGMTNA  474
              V+ S      G  LG  L     A  P VTLGQGYGMT A
Sbjct  303  IRVLKSG-----GAPLGKELEDTVRAKFPNVTLGQGYGMTEA  339
```

- **Nucleotide sequence of fragment 3:**

```
TTGCCGTATTCGTCGGGGTTTACGGGGCTCCCCAAGGGAGTCATTCTAACACACAAGAACTTGGTCACCAGCGTGGCC
CAGCAGGTGGACGGAGAGAATCCAAACCTCTACTTGAAGGAGGACGACGTCGTATTGTGCGTGCTGCCGTTGTTTCAC
ATATTTTCGTTCAACAGCGTGCTGGCTGTGCTCGCTGCGAGCAGGGGCCGGAGTTCTGCTGATGCACAAGTTTGAGAT
AGGTACGCTTCTGGAGCTCATTCAGCGGTACCGAGTGTCGGTGGCAGCGGTGGTGCCGCCGCTGGTTATAGCGCTGGC
GAAGAACCCAATGGTGGCGGAGTTCGATCTGAGCTCTATTAGGGTGGTGTTGTCGGGAGCGGCGCCTCTGGGGAAGGA
GCTGGAGGAGGCGCTCAAGAGCCGAGTCCCTCAGGCAGTGTTGGGTCAGGGCTACGGCATGCCCGAGGCCGAATCACT
AGT
```

- **Amino acid sequence and data bank comparison of fragment 3:**

gi|14289346|gb|AAK58909.1| **4-coumarate:CoA ligase 4** [***Populus balsamifera subsp. trichocarpa x Populus deltoides***] Length=579

Identities = **(78%)** Frame = +2

```
Query  1    LPYSSGFTGLPKGVILTHKNLVTSVAQQVDGENPNLYLKEDDVVLCVLPLFHIFSFNSVL  184
            LP+SSG TGLPKGVILTHK+L+TSVAQQVDGE PNLYLK+DDVVLCVLPLFHIFS NSVL
Sbjct  213  LPFSSGTTGLPKGVILTHKSLITSVAQQVDGEIPNLYLKQDDVVLCVLPLFHIFSLNSVL  272

Query  185  --AVLAASRGRSSADAQV*GTLLELIQRYRVSVAAVVPPLVIALAKNPMVAEFDLSSIRV  358
               ++ A S         ++ G+LLELIQ++ VSVAAVVPPLV+ALAKNPMVA FDLSSIRV
Sbjct  273  LCSLRAGSAVLLMQKFEI-GSLLELIQKHNVSVAAVVPPLVLALAKNPMVANFDLSSIRV  331

Query  359  VLSGAAPLGKELEEALKSRVPQAVLGQGYGMPEA  460
            VLSGAAPLGKELEEAL+SRVPQA+LGQGYGM EA
Sbjct  332  VLSGAAPLGKELEEALRSRVPQAILGQGYGMTEA  365
```

The three cDNA core fragments amplified with the CoA ligase reverse and forward 1 primers were aligned as follows and their identity was calculated.

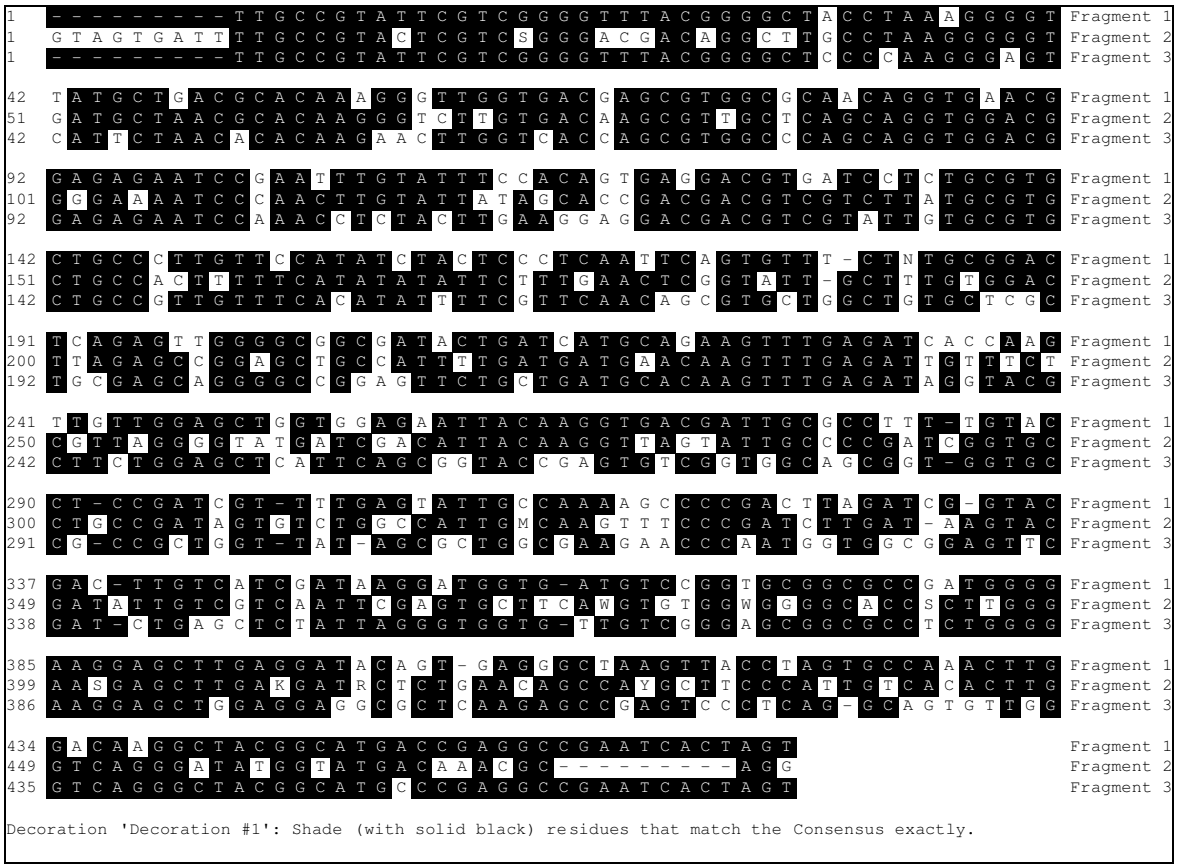


Fig. 4-4: Alignment of the *S. aucuparia* cDNA core fragments 1, 2, and 3 amplified with the CoA ligase reverse and forward 1 primers

Tab. 4-2: Percent identity of the *S. aucuparia* cDNA core fragments 1, 2, and 3 amplified with the CoA ligase reverse and forward 1 primers

Fragment comparison	Percent identity
Fragment 1 with Fragment 2	54.7
Fragment 1 with Fragment 3	60
Fragment 2 with Fragment 3	48

4.4.1.2 Core Fragments Amplified with the CoA Ligase Reverse and Forward 2 Primers

The following three fragments were obtained using this primer pair and turned out to be identical to the core fragments amplified with the CoA ligase reverse and forward 1 primers.

- **Nucleotide sequence of fragment 1:**

```
TTGCCGTATTCGTCGGGGATTACGGGGCTACCTAAAGGGGTTATGCTGACGCACAAAGGGTTGGTGACGAGCGTGGCG
CAACAGGTGGACGGAGAGAATCCGAATTTGTATTTCCACAGTGAGGACGTGATCCTCTGCGTGCTGCCCTTGTTCCAT
ATCTACTCCCTCAATTCAGTGTTTCTCTGCGGACTCAGAGTTGGGGCGGCGATACTGATCGTGCAGAAGTTTGAGATC
ACCAAGTTGTTGGAGCTGGTGGAGAATTACAAGGTGACGATTGCGCCTTTTGTACCTCCGATCGTTTTGAGTATTGCC
AAAAGCCCCGACTTAGATCGGTACGACTTGTCATCGATAAGGATGGTGATGTCCGGTGCGGCGCCGATGGGGAAGGAG
CTTGAGGATACAGTGAGGGCTAAGTTACCTAGTGCCAAACTTGGACARGGCTAYGGCATGACCGAGGCCGAATCTGTG
CTGTCAATGTGCTTAGCATTTGCAAAGGAACCATTTGAGATAAAATCAGGTGCGTGCGGGACTGTTGTAAGAAATGCA
GAGATGAAAATTGTTGACCCTGATACGGGTGCTTCGCTTCCGCGAAATCAAGCTGGAGAGRTTTGCATCAGAGGT
```

- **Amino acid sequence and data bank comparison of fragment 1:**

gi|9651915|gb|AAF91309.1| **4-coumarate:coA ligase 2 [*Rubus idaeus*]**
Length=544

```
            Identities = (91%)        Frame = +1

Query  1    LPYSSGITGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHSEDVILCVLPLFHIYSLNSVF  180
            LPYSSG TGLPKGVMLTHKGLVTSVAQQVDGENPNLYFH EDVILCVLPLFHIYSLNSVF
Sbjct  188  LPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHKEDVILCVLPLFHIYSLNSVF  247

Query  181  LCGLRVGAAILIVQKFEITKLLELVENYKVTIAPFVPPIVLSIAKSPDLDRYDLSSIRMV  360
            LCGLRVGAAILI+QKFEI KLLELVE  KVTIAPFVPPIVLSIAK PDL RYDLSSIRMV
Sbjct  248  LCGLRVGAAILIMQKFEINKLLELVEKEKVTIAPFVPPIVLSIAKCPDLHRYDLSSIRMV  307

Query  361  MSGAAPMGKELEDTVRAKLPSAKLGQGYGMTEAESVLSMCLAFAKEPFEIKSGACGTVVR  540
            MSGAAPMGKELEDTVRAKLP+AKLGQGYGMTEA  VLSMCLAFAKEP+EIKSGACGTVVR
Sbjct  308  MSGAAPMGKELEDTVRAKLPNAKLGQGYGMTEAGPVLSMCLAFAKEPYEIKSGACGTVVR  367

Query  541  NAEMKIVDPDTGASLPRNQAGEICIRG  621
            NAEMKI+DPDT  SLPRNQ+GEICIRG
Sbjct  368  NAEMKIIDPDTNESLPRNQSGEICIRG  394
```

- **Nucleotide sequence of fragment 2:**

```
TTGCCGTACTCGTCCGGGACGACAGGCTTGCCTAAGGGGGTGATGCTAACGCACAAGGGTCTTGTGACAAGCGTTGCT
CAGCAGGTGGACGGGGAAAATCCCAACTTGTATTATAGCACCGACGACGTCGTCTTATGCGTGCTGCCACTTTTTCAT
ATATATTCTTTGAACTCGGTATTGCTTTGTGGACTTAGAGCCGGAGCTGCCATTTTGATGATGAACAAGTTTGAGATT
GTTTCTCKKTAGGGGTATGATCGACATTACAAGGTTAGTATTGCCCCGATCGGTGCCTGCCGATAGTGTCTGGCCATT
GACAAGTTTCCCGATCTTGATAAGTACGATATTGTCGTCAATTCGAGTGCTTCATGTGTGGTGGGGCACCGCTTGGGA
ACGAGCTTGATGATGCTCTGAACAGCCAYGCTTCCCATTGTCACACTTGGTCAGGGATATGGTATGACAAACGCAGGG
CCAGCTTTGACCATGTCATTGTCATTTGCATTGCGACCCTTCGAGGTCTAACCATGTGGATGTGGCACCGTACTCCGA
TATGCATAGCTCAAAATCGTTGATGCTGAATCTGGACTTCTTGTGCCTCGCAACCAGCCTGGTGAAGATTGCATTACA
GGTG
```

- **Amino acid sequence and data bank comparison of fragment 2:**

gi|9651917|gb|AAF91310.1| **4-coumarate:coA ligase 1 [*Rubus idaeus*]**
Length=543

```
            Identities = (69%)  Frame = +1

Query  1    LPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYYSTDDVVLCVLPLFHIYSLNSVL  180
            LPYSSGTTGLPKGVMLTHKGLVTSV+QQVDGENPNLYYS+DDVVLCVLPLFHIYSLNSVL
Sbjct  187  LPYSSGTTGLPKGVMLTHKGLVTSVSQQVDGENPNLYYSSDDVVLCVLPLFHIYSLNSVL  246

Query  181  LCGLRAGAAILMMNKFEIVSX*GYDRHYKVSIAPIGACR*CLAIDKFPDLDKYDI----V  348
            LCGLRAGAAIL+M KFEIVS     + ++VS+API      LAI KFPDLDKYD+    V
Sbjct  247  LCGLRAGAAILLMQKFEIVSLLELMQKHRVSVAPIVPPT-VLAIAKFPDLDKYDLGSIRV  305

Query  349  VNSSASCVVGHRLGTSLMML*TAXLPIVTLGQGYGMTNAGPALTMSLSFALRPFEV*PCG  528
            + S      G  LG  L     A  P VTLGQGYGMT AGP LTMSL+FA  PFEV P G
Sbjct  306  LKSG-----GAPLGKELEDTVRAKFPNVTLGQGYGMTEAGPVLTMSLAFAKEPFEVKPGG  360

Query  529  CGTVLRYA*LKIVDAESGLLVPRNQPGEDCITG  627
            CGTV+R A LKIVD E+G  +PRN PGE CI G
Sbjct  361  CGTVVRNAELKIVDPETGASLPRNHPGEICIRG  393
```

- **Nucleotide sequence of fragment 3:**

```
CTACCGTTCTCKTCGGGGACGACGGGGCTCCCCAAGGGAGTCATTCTAACACACAAGAACTTGGTCACCAGCGTGGCC
CAGCAGGTGGACGGAGAGAATCCAAACCTCTACTTGAAGGAGGACGACGTCGTATTGTGCGTGCTGCCGTTGTTTCAC
ATATTTTCGTTCAACAGCGTGCTRCTGTGCTCGCTGCGAGCAGGGGCCGGAGTTCTGCTGATGCACAAGTTTGAGATA
GGTACGCTTCTGGAGCTCATTCAGCGGTACCGAGTGTCGGTGGCAGCGGTGGTGCCGCCGCTGGTTATAGCGCTGGCG
AAGAACCCAATGGTGGCGGAGTTCGACCTGAGCTCTATTAGGGTGGTGTTGTCGGGAGCGGCGCCTCYGGGGAAGGAG
CTGGAGGAGGCGCTCAAGAGCCGAGTCCCTCAGGCAGTGTTGGGTCAGGGTTATGGGATGACGGAGGCAGGGCCGGTG
CTGTCAATGTGCATGGCATTCGCAAAGGAACCGATGCCAACCAAGTCAGGGTCGTGTGGGACGGTGGTCCGAAATGCA
GAGCTCAAGGTCCTTGACCCTGAAACTGGTCTGTCACTCGGCTATAACCAACCAGGCGAGATTTGCATCCGTGGCT
```

- **Amino acid sequence and data bank comparison of fragment 3:**

gi|14289346|gb|AAK58909.1| **4-coumarate:CoA ligase 4 [*Populus balsamifera subsp. trichocarpa x Populus deltoides*] Length=579**

```
             Identities = (88%)   Frame = +1
Query  1     LPFSSGTTGLPKGVILTHKNLVTSVAQQVDGENPNLYLKEDDVVLCVLPLFHIFSFNSVL  180
             LPFSSGTTGLPKGVILTHK+L+TSVAQQVDGE PNLYLK+DDVVLCVLPLFHIFS NSVL
Sbjct  213   LPFSSGTTGLPKGVILTHKSLITSVAQQVDGEIPNLYLKQDDVVLCVLPLFHIFSLNSVL  272

Query  181   LCSLRAGAGVLLMHKFEIGTLLELIQRYRVSVAAVVPPLVIALAKNPMVAEFDLSSIRVV  360
             LCSLRAG+ VLLM KFEIG+LLELIQ++ VSVAAVVPPLV+ALAKNPMVA FDLSSIRVV
Sbjct  273   LCSLRAGSAVLLMQKFEIGSLLELIQKHNVSVAAVVPPLVLALAKNPMVANFDLSSIRVV  332

Query  361   LSGAAPXGKELEEALKSRVPQAVLGQGYGMTEAGPVLSMCMAFAKEPMPTKSGSCGTVVR  540
             LSGAAP GKELEEAL+SRVPQA+LGQGYGMTEAGPVLSMC+AF+K+P+PTKSGSCGTVVR
Sbjct  333   LSGAAPLGKELEEALRSRVPQAILGQGYGMTEAGPVLSMCLAFSKQPLPTKSGSCGTVVR  392

Query  541   NAELKVLDPETGLSLGYNQPGEICIRG  621
             NAELKV+DPETG SLG NQPGEICIRG
Sbjct  393   NAELKVIDPETGSSLGRNQPGEICIRG  419
```

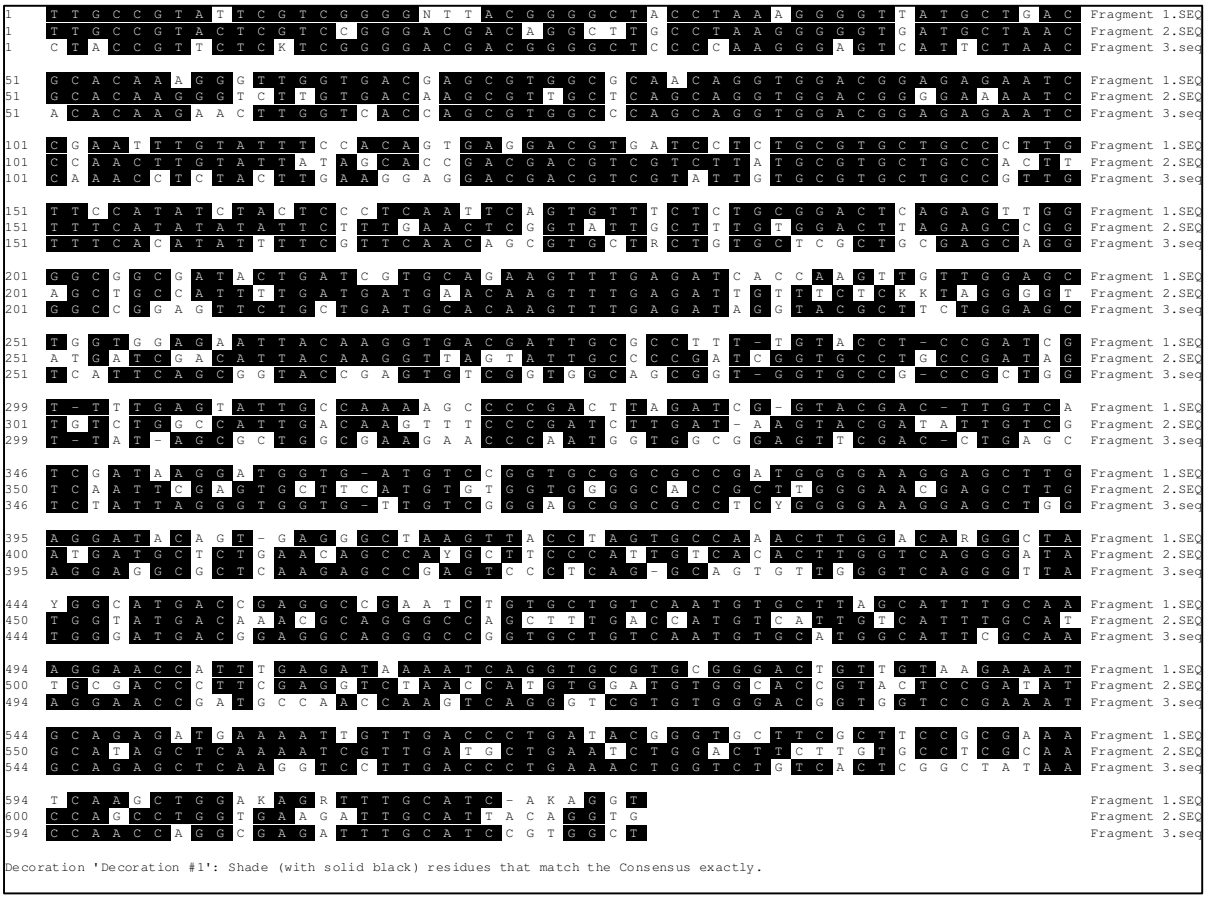


Fig. 4-5: Alignment of the *S. aucuparia* cDNA core fragments 1, 2, and 3 amplified with the reverse and forward 2 primers

Tab. 4-3: Percent identity of the *S. aucuparia* cDNA core fragments 1, 2, and 3 amplified with the reverse and forward 2 primers

Fragment comparison	Percent identity
Fragment 1 with Fragment 2	56
Fragment 1 with Fragment 3	58.8
Fragment 2 with Fragment 3	52.1

4.4.2 5`and 3`Rapid Amplification of cDNA Ends (RACE-PCR)

Based on the core fragments obtained with the degenerate primers, gene specific primers (GSP) were generated (2.8.5). After reverse transcription of *S. aucuparia* cell culture RNA using the gene specific oligonucleotides, cDNA was generated using the Clontech SMART RACE cDNA Amplification Kit and the RNA Ligase Mediated Rapid Amplification of cDNA Ends Kit (RLM-RACE) from Ambion as described in 3.2.1 and 3.2.3.

The 5`RACE Ready cDNA was used in a PCR with gene specific reverse primers and the anchor primer long universal primer ( $RACE_{long}$ or $RACE_{long+short}$ ), whereas the 3`RACE PCR was performed with a gene specific forward primer and either Oligo-dT or $RACE_{long}$ primers as the reverse primer for SMART RACE in a PCR containing 3`cDNA (Fig 4-6 and 4-7).

In RLM-RACE, 5' and 3' RACE primers (outer and inner primers) are complimentary to the anchored adapter. The 3' RACE outer primer was used in PCR with a gene specific primer to amplify 3' cDNA product, while the 5' RACE outer primer was used in PCR to amplify 5' cDNA product. Nested PCR was performed using the 5' RACE inner primer (Fig.4-9).

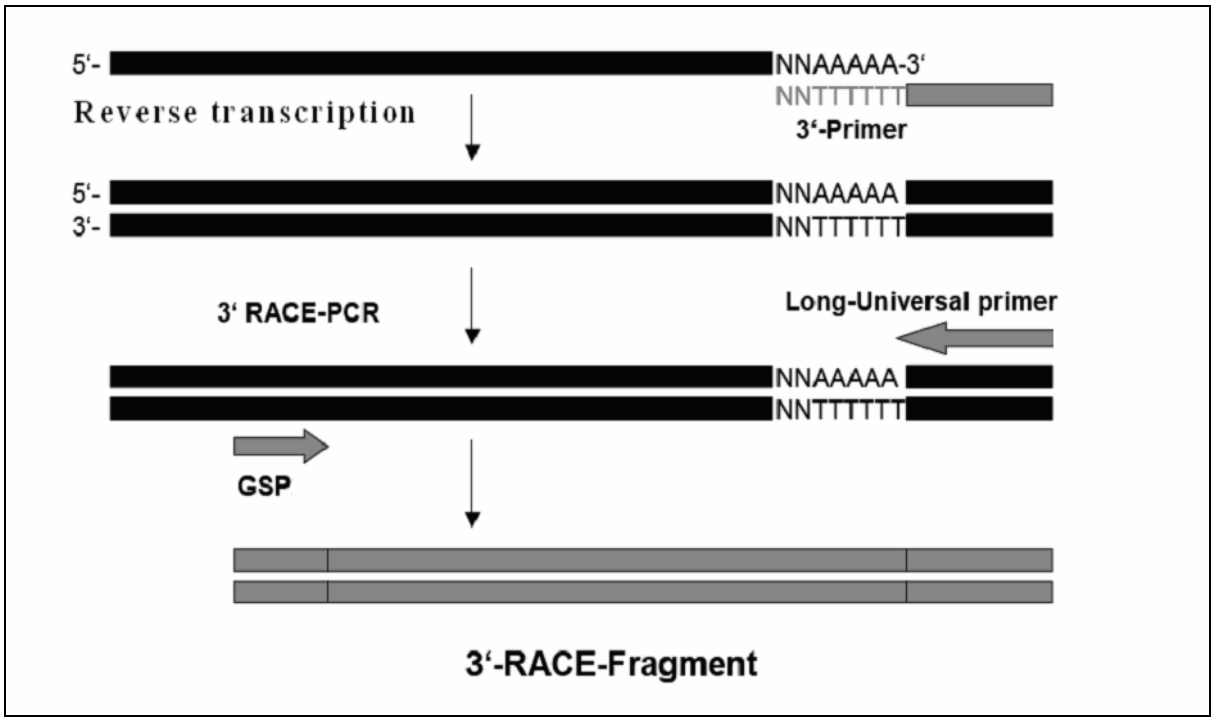


Fig. 4-6: Mechanism of SMART cDNA synthesis (3´- RACE)

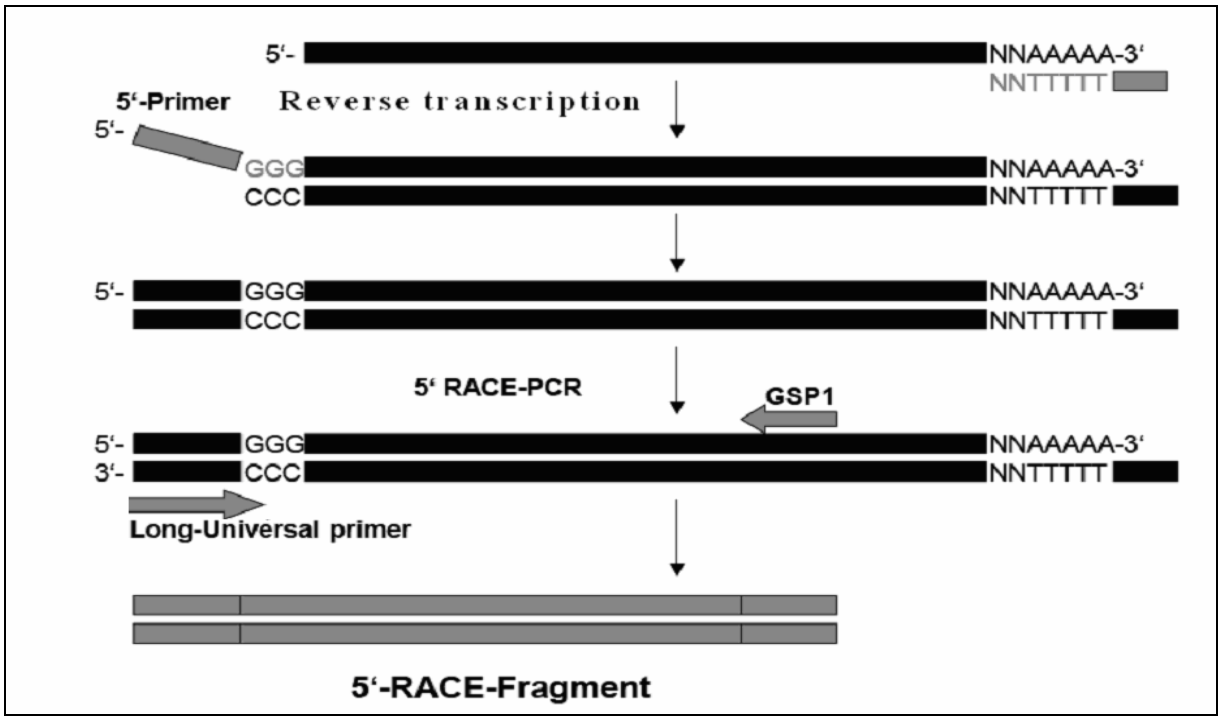


Fig. 4-7: Mechanism of SMART cDNA synthesis (5´- RACE)

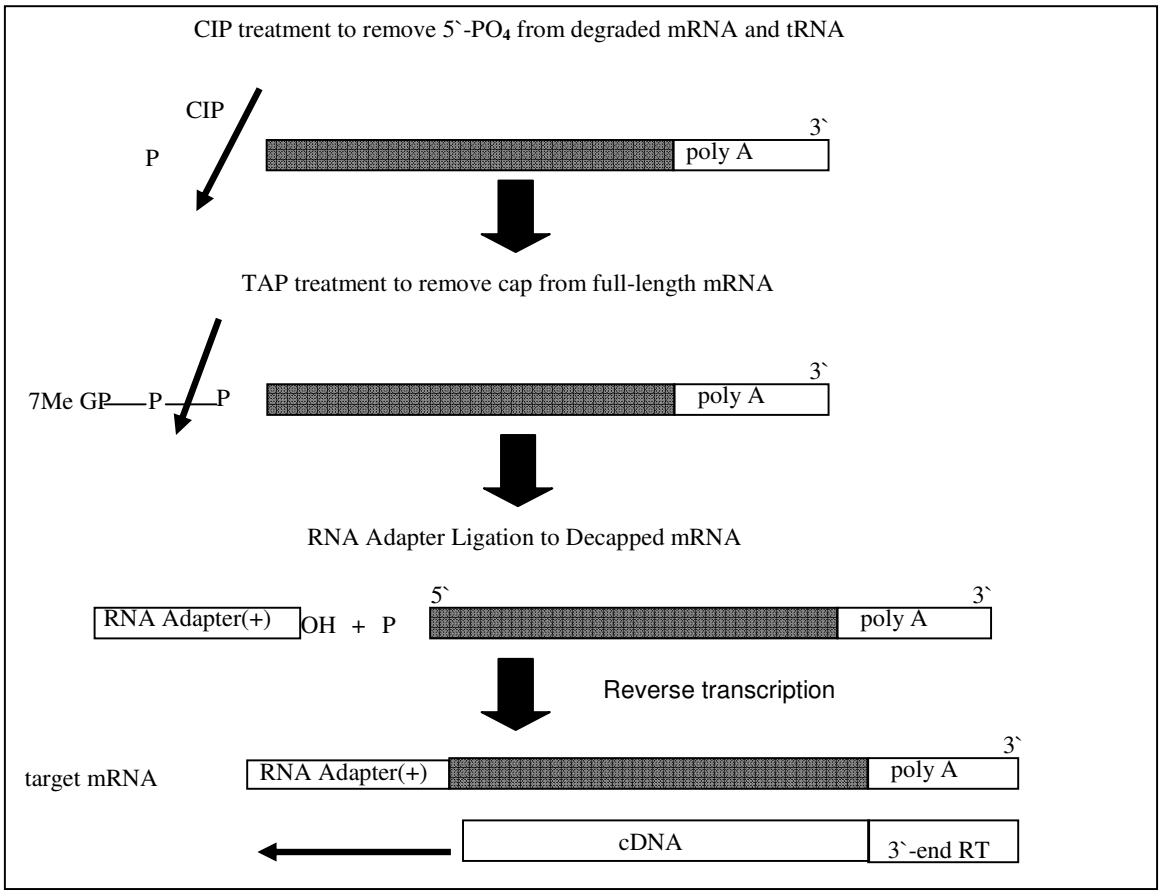


Fig. 4-8: Overview of the First choice RLM-RACE Kit

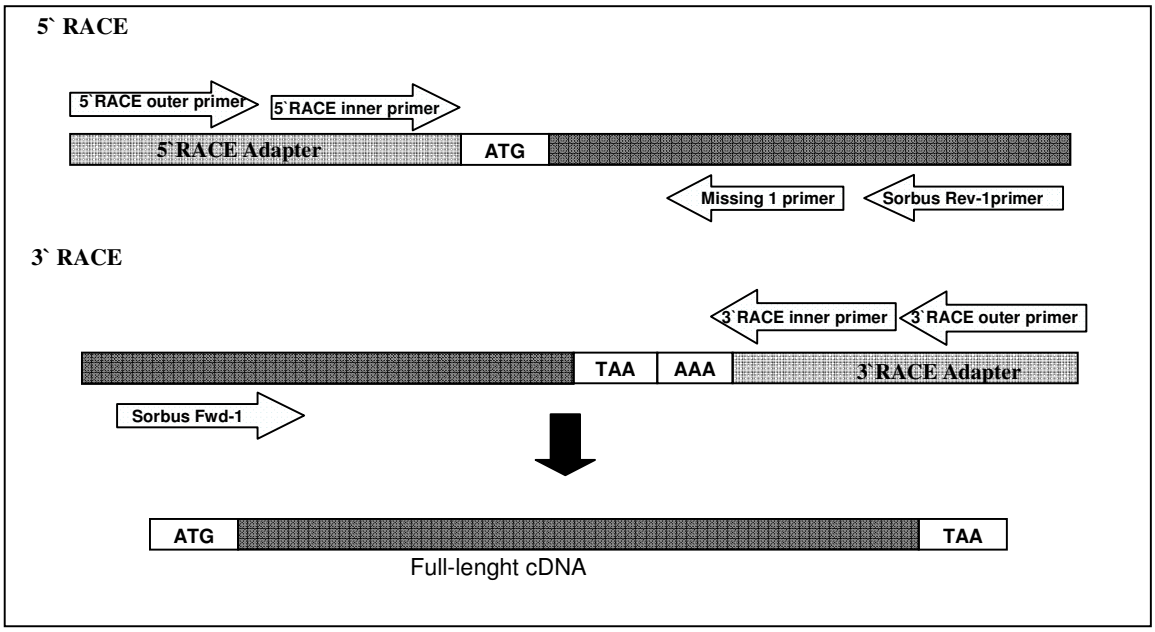


Fig. 4-9: Overview of the First choice RLM-RACE Kit (5`and 3`RACE)

To deduce full-length sequences for the partial coenzyme A ligase cDNAs, a “Touch Down” temperature program was used, as given in the following table:

Tab. 4-4: Temperature program for 5`-RACE-PCR

Step	Temp. (°C)	Time (Sec)	Cycle	Δ Temp (°C)	To step
1	95	180			
2	70	60			
3	94	30			
4	60	40		-1	
5	72	90	10		3
6	94	30			
7	50	40			
8	72	90	20		6
9	72	720			

4.4.3 Isolation and Sequence Analysis of cDNA Fragments obtained with Gene Specific Primers

Amplified cDNA fragments from the 5’and 3` RACE-PCR were cloned in the pGEM®-T vector and transformed (3.9.1). Plasmid DNA was isolated and subjected to restriction enzyme digestion (3.10.3.3). Each reaction was then loaded into a 1% agarose gel in TAE to verify the presence of inserts. The resulting different cDNA fragments were cloned and sequenced. In all cases, they were shown to be identical in the overlapping portions when compared to the respective partial cDNAs. One composed sequence contained an open reading frame (ORF) predicted to encode a polypeptide of 605 amino acids.

The following 3`and 5 `-RACE PCR fragments were obtained:

Fragment 1:

This fragment was obtained with gene specific primers deduced from the fragment 1 obtained under 4.4.1.

- **Nucleotide sequence of fragment 1:**

```
TTGCCGTATTCGTCGGGGNTTACGGGGCTACCTAAAGGGGTTATGCTGACGCACAAAGGGTTGGTGACGAGCGTGGCG
CAACAGGTGGACGGAGAGAATCCGAATTTGTATTTCCACAGTGAGGACGTGATCCTCTGCGTGCTGCCCTTGTTCCAT
ATCTACTCCCTCAATTCAGTGTTTCTCTGCGGACTCAGAGTTGGGGCGGCGATACTGATCGTGCAGAAGTTTGAGATC
ACCAAGTTGTTGGAGCTGGTGGAGAATTACAAGGTGACGATTGCGCCTTTTGTACCTCCGATCGTTTTGAGTATTGCC
AAAAGCCCCGACTTAGATCGGTACGACTTGTCATCGATAAGGATGGTGATGTCCGGTGCGGCGCCGATGGGGAAGGAG
CTTGAGGATACAGTGAGGGCTAAGTTACCTAGTGCCAAACTTGGACARGGCTAYGGCATGACCGAGGCCGAATCTGTG
CTGTCAATGTGCTTAGCATTTGCAAAGGAACCATTTGAGATAAAATCAGGTGCGTGCGGGACTGTTGTAAGAAATGCA
GAGATGAAAATTGTTGACCCTGATACGGGTGCTTCGCTTCCGCGAAATCAAGCTGGAKAGRTTTGCATCAKAGGTAGC
CAAATCATGAARGGTACACAGAGTGGCTACTATTCCATTACAATAAACCAATAATTTCCTCAAAAAAAAAAAAAAAAA
AAAAAAAAAAA
```

- **Amino acid sequence and data bank comparison of fragment 1:**

gi|9651915|gb|AAF91309.1| **4-coumarate:coA ligase 2 [*Rubus idaeus*]**

```
Length=544
          Identities = (91%)                    Frame = +1

 Query   1 LPYSSGXTGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHSEDVILCVLPLFHIYSLNSVF  180
           LPYSSG TGLPKGVMLTHKGLVTSVAQQVDGENPNLYFH EDVILCVLPLFHIYSLNSVF
Sbjct 188 LPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHKEDVILCVLPLFHIYSLNSVF  247

Query 181 LCGLRVGAAILIVQKFEITKLLELVENYKVTIAPFVPPIVLSIAKSPDLDRYDLSSIRMV  360
           LCGLRVGAAILI+QKFEI KLLELVE  KVTIAPFVPPIVLSIAK PDL RYDLSSIRMV
Sbjct 248 LCGLRVGAAILIMQKFEINKLLELVEKEKVTIAPFVPPIVLSIAKCPDLHRYDLSSIRMV  307

Query 361 MSGAAPMGKELEDTVRAKLPSAKLGQGYGMTEAESVLSMCLAFAKEPFEIKSGACGTVVR  540
           MSGAAPMGKELEDTVRAKLP+AKLGQGYGMTEA  VLSMCLAFAKEP+EIKSGACGTVVR
Sbjct 308 MSGAAPMGKELEDTVRAKLPNAKLGQGYGMTEAGPVLSMCLAFAKEPYEIKSGACGTVVR  367

Query 541 NAEMKIVDPDTGASLPRNQAGXXCIXGSQIMKG  639
           NAEMKI+DPDT  SLPRNQ+G  CI GSQIMKG
Sbjct 368 NAEMKIIDPDTNESLPRNQSGEICIRGSQIMKG  400
```

Fragment 2:

This fragment was obtained with gene specific primers deduced from the fragment 2 obtained under 4.4.1.

- **Nucleotide sequence of fragment 2:**

```
GTCCACAACAAAAACCACTCGAGCTCTAAGCCCTGTATCATCGACGGCGCCACCGGAGACATATACACCTTCGCCGAT
GTGGAACTTAACGCGCGCAGAGTCGCGTCAGGKCTCAACAAGCTTGGAATCCAACAAGGCGACGTCATCATGCTCTTG
CTCCCCAACTCCCCAGCGTTCGCCTTCGCCTTCCTCGGAGCCTCCTTTGGGGGCGCTATGTCCATCCGCGGCSGACCC
CTTTCTTCCACACCCGCCGAGATCTTTAAACCAGGCCAAGGGCCTCAAAAGTTAAACTCATCATCACCTTAGCGTGCT
CTTASGACAAGGGTCAAGGACTTATCGTCATCAAGTGACGATGGTTCACGACCATTAAGCTAATGGTGGGTCGACTTC
TCCGCCTGATCCGAAGCTGTTTGCATTTCTCCGAGCTTCTTCAAGCCGATGAAAACGACATGCCGGAGGTTGACATCA
GCCCAGACGACGTAGTGATTTTGCCGTACTCGTCCGGGACGACAGGCTTGCCTAAGGGGGTGATGCTAACGCACAAGG
GTCTTGTGACAAGCGTTGCTCAGCAGGTGGACGGGGAAAATCCCAACTTGTATTATAGCACCGACGACGTCGTCTTAT
GCGTGCTGCCACTTTTTCATATATATTCTTTGAACTCGGTATTGCTTTGTGGACTTAGAGCCGGAGCTGCCATTTTGA
TGATGAACAAGTTTGAGATTGTTTCTCKKTAGGGGTATGATCGACATTACAAGGTTAGTATTGCCCCGATCGgTGCCT
GCCGATAGTGTCTGGCCATTGACAAGTTTCCCGATCTTGATAAGTACGATATTGTCGTCAATTCGAGTGCTTCATGTG
TGGTGGGGCACCGCTTGGGAACGAGCTTGATGATGCTCtGAACAGCCAYGCTTCCCATTGTCACACTTGGTCAGGGAT
ATGGTATGACAAACGCAGGGCCAGCTTTGACCATGTCATTGTCATTTGCATTGCGACCCTTCGAGGTCTAACCATGTG
GATGTGGCACCGTACTCCGATATGCATAGCTCAAAATCGTTGATGCTGAATCTGGACTTCTTGTGCCTCGCAACCAGC
CTGGTGAAGATTGCATTACAGGTGACCAGATCATGACAGGTCATCTTAATGATCCGGAGTCTTCATTGGCACCCATAG
ACCATGAAAGCGGGCTACACGCCGGTGATATAAGCTTCATTGATGATGATGATGACTATTCATTGTTTGATCGGTTGA
CGGAACTGTTCACTACCAAGGATTTCAAGTGGCCCCTGCTGAACTTTGAAGCCTG
```

- **Amino acid sequence and data bank comparison of fragment 2:**

gi|9651917|gb|AAF91310.1| **4-coumarate:coA ligase 1** [***Rubus idaeus***]
Length=543

```
             Identities = (66%)  Frame = +1

Query  1    NKNHSSSKPCIIDGATGDIYTFADVELNARRVASGLNKLGIQQGDVIMLLLPNSPAFAFA  186
            NK+H +SKPCII+G TGDI+T+A  +L AR+VASGLNKLGI++GDV MLLLPN+  F FA
Sbjct  34   NKSHLTSKPCIINGTTGDIHTYAKFKLTARKVASGLNKLGIEKGDVFMLLLPNTSEFVFA  93

Query  187  FLGASFGGAMSIRGXPLSSTPAEIFK  264
            FLGASF GAM     P   TPAEI K
Sbjct  94   FLGASFCGAMMTAANPF-FTPAEIAK  118
             Identities = (61%), Frame = +3

Query  234  PFFHTRRDL*TRPRASKVKLIITLACSXDKGQGLIVIK*RWFTTIKLMVGRLLRLIRSCL  413
            PFF T  ++  + +ASK KLIIT AC  DK + L   +      +KLM         SCL
Sbjct  109  PFF-TPAEIAKQAKASKAKLIITFACYYDKVKDLSCDE------VKLMCIDSPPPDSSCL  161

Query  414  HFSELLQADENDMPEVDISPDDVVILPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPN  593
            HFSEL Q+DEND+P+VDISPDDVV LPYSSGTTGLPKGVMLTHKGLVTSV+QQVDGENPN
Sbjct  162  HFSELTQSDENDVPDVDISPDDVVALPYSSGTTGLPKGVMLTHKGLVTSVSQQVDGENPN  221

Query  594  LYYSTDDVVLCVLPLFHIYSLNSVLLCGLRAGAAILMMNKFEIVSX*GYDRHYKVSIAPI  773
            LYYS+DDVVLCVLPLFHIYSLNSVLLCGLRAGAAIL+M KFEIVS     + ++VS+API
Sbjct  222  LYYSSDDVVLCVLPLFHIYSLNSVLLCGLRAGAAILLMQKFEIVSLLELMQKHRVSVAPI  281

Query  774  GACR*CLAIDKFPDLDKYDI----VVNSSASCVVGHRLGTSLMML*TAXLPIVTLGQGYG  941
                  LAI KFPDLDKYD+    V+ S      G  LG  L     A  P VTLGQGYG
Sbjct  282  VPPT-VLAIAKFPDLDKYDLGSIRVLKSG-----GAPLGKELEDTVRAKFPNVTLGQGYG  335

Query  942  MTNAGPALTMSLSFALRPFEV*PCGCGTVLRYA*LKIVDAESGLLVPRNQPGEDCITGDQ  1121
            MT AGP LTMSL+FA  PFEV P GCGTV+R A LKIVD E+G  +PRN PGE CI G Q
Sbjct  336  MTEAGPVLTMSLAFAKEPFEVKPGGCGTVVRNAELKIVDPETGASLPRNHPGEICIRGHQ  395

Query  1122 IMTGHLNDPESSLAPIDHESGLHAGDISFIDDDDDYSLFDRLTELFTTKDFKWPLLNFEA  1301
            IM G+LNDPE++   ID +  LH GDI FIDDD++  + DRL EL   K F+      EA
Sbjct  396  IMKGYLNDPEATRTTIDKQGWLHTGDIGFIDDDEELFIVDRLKELIKYKGFQVAPAELEA  455
```

Fragment 3:

This fragment was obtained with gene specific primers deduced from the fragment 3 obtained under 4.4.1.

- **Nucleotide sequence of fragment 3 :**

```
CTACCGTTCTCKTCGGGGACGACGGGGCTCCCCAAGGGAGTCATTCTAACACACAAGAACTTGGTCACCAGCGTGGCC
CAGCAGGTGGACGGAGAGAATCCAAACCTCTACTTGAAGGAGGACGACGTCGTATTGTGCGTGCTGCCGTTGTTTCAC
ATATTTTCGTTCAACAGCGTGCTRCTGTGCTCGCTGCGAGCAGGGGCCGGAGTTCTGCTGATGCACAAGTTTGAGATA
GGTACGCTTCTGGAGCTCATTCAGCGGTACCGAGTGTCGGTGGCAGCGGTGGTGCCGCCGCTGGTTATAGCGCTGGCG
AAGAACCCAATGGTGGCGGAGTTCGACCTGAGCTCTATTAGGGTGGTGTTGTCGGGAGCGGCGCCTCYGGGGAAGGAG
CTGGAGGAGGCGCTCAAGAGCCGAGTCCCTCAGGCAGTGTTGGGTCAGGGTTATGGGATGACGGAGGCAGGGCCGGTG
CTGTCAATGTGCATGGCATTCGCAAAGGAACCGATGCCAACCAAGTCAGGGTCGTGTGGGACGGTGGTCCGAAATGCA
GAGCTCAAGGTCCTTGACCCTGAAACTGGTCTGTCACTCGGCTATAACCAACCAGGCGAGATTTGCATCCGTGGCTTT
CAAATTATGAAAGGATATTTGAATGACGCTGCGGCTACGGCAACAACCTTAGACACGGAGGGCTGGCTTCACACTGGT
GACGTGGGTTATGTGGATGATGACGATGAGGTTTTCATCGTTGACAGAGTCAAGGAGCTCATCAAATTCAAAGGCTTC
CAAGTGCCACCAKCTGAKCTGGAGTCCCTCCTTATAAGCCACCCATCAATTGCAGATGCGGCCGTTGTTCCGCAAAGA
GATGATGCTGCTGGTGAGGTTCCCGTTGCATTTGTGGTTCGGTCTGATGGTCTTGAACTTACYGAAGAGGCTGTAAAA
GAATTTATAGCAAAACAGGTGGTGTTTTACAAGAGACTGCACAAGGTGCACTTTGTCCATGCAATTCCAAAGTCTGCG
TCTGGAAAGATCTTGAGAAAAGACCTGAGAGCCAAGCTTGCCACCGCAACCCCACCGCTGCCCTAATTTTGATGAATC
ATCTTAATTAGCATGTACTGTATCATAATGTTATTATTTATCCCATTTGAAGGTGAAAACCGGGCTCGATCAATAATA
CGTGTAATTCTTTTGAAACACTATATCAACTTTGGTATCCATCATTTCATATCATATCAAATTTGTTGATGAATAAGA
ACATGACGATTTTAATTGGTACCTTTAAAAAAAAAAAAAAAAAAAAA
```

- **Amino acid sequence and data bank comparison of fragment 3:**

gi|9651913|gb|AAF91308.1| **4-coumarate:coA ligase 3 [*Rubus idaeus*]**

```
Length=591
             Identities = (83%)    Frame = +1
Query  1     LPFSSGTTGLPKGVILTHKNLVTSVAQQVDGENPNLYLKEDDVVLCVLPLFHIFSFNSVL  180
             LPFSSGTTGLPKGVILTHK+L+TSVAQQVDGENPNLYLK DDVVLCVLPLFHIFS NSVL
Sbjct  228   LPFSSGTTGLPKGVILTHKSLITSVAQQVDGENPNLYLKGDDVVLCVLPLFHIFSLNSVL  287

Query  181   LCSLRAGAGVLLMHKFEIGTLLELIQRYRVSVAAVVPPLVIALAKNPMVAEFDLSSIRVV  360
             LCSLRAGA VL+M KFEIGTLLELIQRYRV           +  +  MVA++DLSSIRVV
Sbjct  288   LCSLRAGAAVLVMPKFEIGTLLELIQRYRVFCGGGWCLAGDSAGEESMVADYDLSSIRVV  347

Query  361   LSGAAPXGKELEEALKSRVPQAVLGQGYGMTEAGPVLSMCMAFAKEPMPTKSGSCGTVVR  540
             LSGAAP GKELEEAL++RVPQAVLGQGYGMTEAGPVLSMC+AFAK+P PTKSGSCG+VVR
Sbjct  348   LSGAAPLGKELEEALRNRVPQAVLGQGYGMTEAGPVLSMCLAFAKQPFPTKSGSCGSVVR  407

Query  541   NAELKVLDPETGLSLGYNQPGEICIRGFQIMKGYLNDAAATATTXDTEGWLHTGDVGYVD  720
             NAELKV++PETG SLGYNQPGEIC+RG QIMKGYLND  ATATT D EGWLHTGD+GYVD
Sbjct  408   NAELKVVEPETGRSLGYNQPGEICVRGSQIMKGYLNDGEATATTVDVEGWLHTGDIGYVD  467

Query  721   DDDEVFIVDRVKELIKFKGFQVPPXXLESLLISHPSIADAAVVPQRDDAAGEVPVAFVVR  900
             DDDEVFIVDRVKELIKFKGFQVPP  LESLLISHPS+    VVPQ+DDAAGEVPVAFVVR
Sbjct  468   DDDEVFIVDRVKELIKFKGFQVPPAELESLLISHPSMQMQLVVPQKDDAAGEVPVAFVVR  527

Query 901    SD-GLELTEEAVKEFIAKQVVFYKRLHKVHFVHAIPKSASGKILRKDLRAKLATATPPLP 1077
             S+ G ELTEEAVKEFIAKQVVFYKRLHKV+FVHAIPKS +GKILRKDLRAKLA A  P P
Sbjct  528   SNGGNELTEEAVKEFIAKQVVFYKRLHKVYFVHAIPKSPAGKILRKDLRAKLAAAATPNP  587
```

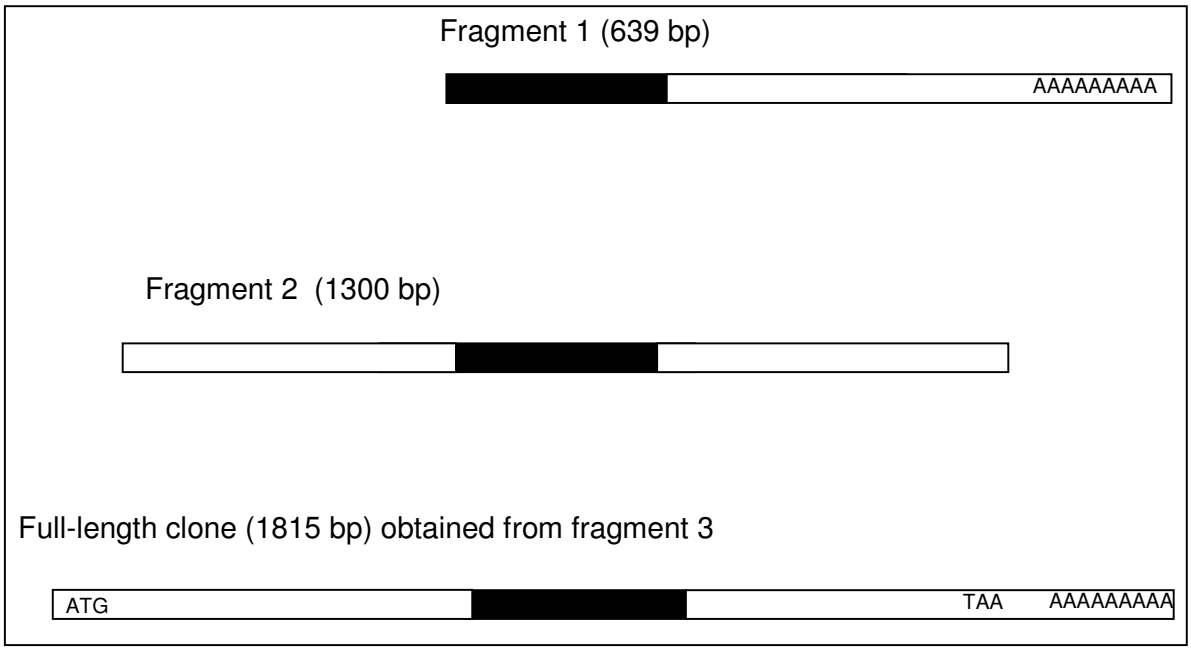


Fig. 4-10: Schematic representation of the three *S. aucuparia* clones. The core fragments obtained with degenerate primers (black boxes) were extended to the 5' and 3' ends as shown. A full- length clone resulted from fragment 3.

During my attempts to get full-length clones with gene specific primers from the core fragments of *S. aucuparia,* additional cDNA fragments have been cloned and sequenced. In all cases, they were shown to be identical in the overlapping portions but different in the extended regions when compared to the above three fragments (Fig. 4-4, 4-5). These nucleotide sequences and amino acid comparisons are found in the appendix.

4.4.4 Full-length Clone from *S. aucuparia* cell cultures

Based on the composed coenzyme A ligase sequence, a new set of PCR primers was designed specific to the 5`and 3`ends of the ORF which was cloned using the RNA Ligase Mediated Rapid Amplification of cDNA Ends kit (RLM-RACE). This allowed the amplification of the full-length coding region of the coenzyme A ligase. The primers were designed as follows: The reverse primer (OVER EXP 3`) contained the EcoRI restriction site integrating the stop codon TAA, whereas the forward primer (OVER EXP 5`) contained a NheI restriction site integrating the start codon ATG (Tab. 4-5).

Tab. 4-5: Gene specific primers used for the amplification of the *S. aucuparia* 4CL ORF

Oligonucleotide Sequence	Tm (°C)
Over. Exp. 3’: EcoRI 5`-**GAA TTC** TTA GGG CAG CGG TGG GGT TGC GGTG-3` Over. Exp. 5’: NheI 5`-**GCTAGC**ATG ATC TCC ATT GCT TCT AAT TCC GTT -3`	73 67

- **Nucleotide sequence of the full length clone**

ATGATCTCCATTGCTTCTAATTCCGTTGAAACCCAAAAGGCGGCAGACACAGCTACCAATCTCATGCCTCCTCTGATT
AATTCTACCTCCCAACAAAACCTAACCAAATTGCAACCCCCCGCCTGCACCAACAATATTATTGATTCCACCACCGCC
ACCGCCACCATTAACCATGTATTCAGATCAAAACTACCAGACATAGCCATCCCCAACCACCTCCCTCTCCACACCTAC
TGCTTTCAAAACCTCCCCGAGTTCTCCGACAGGCCTTGCTTGATCGTGGGCTCCACCGGAAAATCATACTCTTTCTCC
GAGACCCACCTCATTTCTCAGAAGACCGGCGCCGGCCTCTCCAACCTCGGCATCCAAAAAGGCGACGTCATCATGATC
CTCCTCCAAAACTGCGCAGAGTTCGTCTTCGCTTTTATGGGCGCTTCCCTGATCGGCGCCGTTACCACCACCGCCAAC
CCCTTCTACACCACTGCCGAGATTTTCAAGCAGGTCAAGGCCGCTAATGCCAAACTCATCATCACTCAATCCCAGTAC
GTCAATAAGCTCCGAGAACATCCCTCGTCCGCCGACGGTGCCGACCAGAATAACTACCCAAAACTCGGCGAAGACTTT
AAGGTCGTCACAATTGACGATCCTCCGGAGAATTGCTTGCATTTCTCCGTGCTCTCCGAGGCCAACGAGAAGGAGCTT
CCGGACGTGGTGATCGACGCGGAGGACCCGGTGGCCCTACCGTTCTCTTCGGGGACGACGGGGCTCCCCAAGGGAGTC
ATTCTAACACACAAGAACTTGGTCACCAGCGTGGCCCAGCAGGTGGACGGAGAGAATCCAAACCTCTACTTGAAGGAG
GACGACGTCGTATTGTGCGTGCTGCCGTTGTTTCACATATTTTCGTTCAACAGCGTGCTGCTGTGCTCGCTGCGAGCA
GGGGCCGGAGTTCTGCTGATGCACAAGTTTGAGATAGGTACGCTTCTGGAGCTCATTCAGCGGTACCGAGTGTCGGTG
GCAGCGGTGGTGCCGCCGCTGGTTATAGCGCTGGCGAAGAACCCAATGGTGGCGGAGTTCGACCTGAGCTCTATTAGG
GTGGTGTTGTCGGGAGCGGCGCCTCTGGGGAAGGAGCTGGAGGAGGCGCTCAAGAGCCGAGTCCCTCAGGCAGTGTTG
GGTCAGGGTTATGGGATGACGGAGGCAGGGCCGGTGCTGTCAATGTGCATGGCATTCGCAAAGGAACCGATGCCAACC
AAGTCAGGGTCGTGTGGGACGGTGGTCCGAAATGCAGAGCTCAAGGTCCTTGACCCTGAAACTGGTCTGTCACTCGGC
TATAACCAACCAGGCGAGATTTGCATCCGTGGCTTTCAAATTATGAAAGGATATTTGAATGACGCTGCGGCTACGGCA
ACAACCATAGACACGGAGGGCTGGCTTCACACTGGTGACGTGGGTTATGTGGATGATGACGATGAGGTTTTCATCGTT
GACAGAGTCAAGGAGCTCATCAAATTCAAAGGCTTCCAAGTGCCACCAGCTGAGCTGGAGTCCCTCCTTATAAGCCAC
CCATCAATTGCAGATGCGGCCGTTGTTCCGCAAAGAGATGATGCTGCTGGTGAGGTTCCCGTTGCATTTGTGGTTCGG
TCTGATGGTCTTGAACTTACTGAAGAGGCTGTAAAAGAATTTATAGCAAAACAGGTGGTGTTTTACAAGAGACTGCAC
AAGGTGCACTTTGTCCATGCAATTCCAAAGTCTGCGTCTGGAAAGATCTTGAGAAAAGACCTGAGAGCCAAGCTTGCC
ACCGCAACCCCACCGCTGCCC**TAA**

- **Amino acid sequence and data bank comparison of the full length clone:**

gi|9651913|gb|AAF91308.1| **4-coumarate:CoA ligase 3 [*Rubus idaeus*]**

```
Length=591

             Identities = (74%) Frame = +1


Query  1     MISIASNS-------VETQKAADTATNLMPPLINSTSQQNLTKLQPPACTNNIIDSTTAT  159
             MISIASN+       VET    + + N++  +I+++  Q   K QPP            T
Sbjct  1     MISIASNNNNNNSVVVETPTKPEISPNIISDVISTSQTQPEQKQQPP------------T  48

Query  160   ATINHVFRSKLPDIA-IPNHLPLHTYCFQNLPEFSDRPCLIVGSTGKSYSFSETHLISQK  336
              T +HVF+SKLPD+  I NHLPLHTYCF+N+ +FS+RPCLI+GSTGKSY+FSET L+SQK
Sbjct  49    TTTHHVFKSKLPDLPNISNHLPLHTYCFENISDFSERPCLIIGSTGKSYTFSETRLLSQK  108

Query  337   TGAGLSNLGIQKGDVIMILLQNCAEFVFAFMGASLIGAVTTTANPFYTTAEIFKQVKAAN  516
             TG GLS LGI KGDV+MILLQNCAEFVFAFMGAS+IGAVTTTANPFYT +EIFKQ++A+N
Sbjct  109   TGVGLSKLGIHKGDVVMILLQNCAEFVFAFMGASMIGAVTTTANPFYTASEIFKQLEASN  168

Query  517   AKLXITQSQYVNKLREHPSSADGADQNNYPKLGEDFKVVTIDDPPENCLHFSVLSEANEK  696
             AKL ITQSQYV+KLR+                G+ F+VVTIDDPPENCLHFSVLS+ANE
Sbjct  169   AKLIITQSQYVDKLRQP---------------GQHFQVVTIDDPPENCLHFSVLSDANEN  213

Query  697   ELPDVVIDAEDPVALPFSSGTTGLPKGVILTHKNLVTSVAQQVDGENPNLYLKEDDVVLC  876
             ELP V ID +DPVALPFSSGTTGLPKGVILTHK+L+TSVAQQVDGENPNLYLK DDVVLC
Sbjct  214   ELPQVSIDPDDPVALPFSSGTTGLPKGVILTHKSLITSVAQQVDGENPNLYLKGDDVVLC  273

Query  877   VLPLFHIFSFNSVLLCSLRAGAGVLLMHKFEIGTLLELIQRYRVSVAAVVPPLVIALAKN 1056
             VLPLFHIFS NSVLLCSLRAGA VL+M KFEIGTLLELIQRYRV           +  +
Sbjct  274   VLPLFHIFSLNSVLLCSLRAGAAVLVMPKFEIGTLLELIQRYRVFCGGGWCLAGDSAGEE  333

Query  1057  PMVAEFDLSSIRVVLSGAAPLGKELEEALKSRVPQAVLGQGYGMTEAGPVLSMCMAFAKE 1236
              MVA++DLSSIRVVLSGAAPLGKELEEAL++RVPQAVLGQGYGMTEAGPVLSMC+AFAK+
Sbjct  334   SMVADYDLSSIRVVLSGAAPLGKELEEALRNRVPQAVLGQGYGMTEAGPVLSMCLAFAKQ  393

Query  1237  PMPTKSGSCGTVVRNAELKVLDPETGLSLGYNQPGEICIRGFQIMKGYLNDAAATATTID 1416
             P PTKSGSCG+VVRNAELKV++PETG SLGYNQPGEIC+RG QIMKGYLND  ATATT+D
Sbjct  394   PFPTKSGSCGSVVRNAELKVVEPETGRSLGYNQPGEICVRGSQIMKGYLNDGEATATTVD  453

Query  1417  TEGWLHTGDVGYVDDDDEVFIVDRVKELIKFKGFQVPPAELESLLISHPSIADAAVVPQR 1596
              EGWLHTGD+GYVDDDDEVFIVDRVKELIKFKGFQVPPAELESLLISHPS+    VVPQ+
Sbjct  454   VEGWLHTGDIGYVDDDDEVFIVDRVKELIKFKGFQVPPAELESLLISHPSMQMQLVVPQK  513

Query  1597  DDAAGEVPVAFVVRSD-GLELTEEAVKEFIAKQVVFYKRLHKVHFVHAIPKSASGKILRK 1773
             DDAAGEVPVAFVVRS+ G ELTEEAVKEFIAKQVVFYKRLHKV+FVHAIPKS +GKILRK
Sbjct  514   DDAAGEVPVAFVVRSNGGNELTEEAVKEFIAKQVVFYKRLHKVYFVHAIPKSPAGKILRK  573

Query  1774  DLRAKLATATPPLP  1815
             DLRAKLA A  P P
Sbjct  574   DLRAKLAAAATPNP  587
```

4.5 Heterologous Expression

A full-length PCR was performed by using pfu-DNA polymerase (Cline et al., 1996) in a standard 25 µl-reaction volume at an annealing temperature of 58°C. ORF was ligated into the NheI/EcoRI-digested pRSET vector containing a polyhistidine-encoding region (3.14.1). The expression vector was transferred into *E. coli* strain BL21 and protein expression was induced by IPTG and the bacterial culture was further incubated at 25°C over night (3.14.3). The efficiency of this procedure was examined by SDS-PAGE of a protein extract from the IPTG-treated cells, resulting in the detection of an overexpressed protein of approximately 70 kDa (Fig. 4-12).

4.6 Enzyme Assays with Crude *E. coli* Protein Extract

To check prior to large-scale cultivation of *E. coli* and protein purification whether the overexpressed protein was functional, the soluble protein fraction from cell-free extracts of IPTG-treated bacteria was subjected to activity tests. Enzyme assays were carried out using three possible phenylpropanoid substrates, namely p-coumaric, cinnamic, and benzoic acids. Product analysis was performed using spectrophotometry and HPLC for the first two substrates and HPLC alone for benzoic acid (3.15.2). CoA ligase activity was observed, with a strong preference for p-coumaric acid and decreasing activities toward cinnamic and benzoic acids. In HPLC analysis, cinnamoyl-, p-coumaroyl-, and benzoyl-CoAs were co-eluted with the enzymatic products.

4.7 Protein Purification

FPLC was used to isolate the recombinant protein from the crude protein extract (3.14.6). The N-terminal 6xHis tag enabled purification on a His-trap column (3.14.5). The elution profile is shown in (Fig. 4-11). Fraction 4 contained the bulk of recombinant protein. The yield of the pure protein, as determined by the Bradford method, was typically 30–50 mg/l of culture. Purified protein could be stored at -20 or -80°C in 100% (v/v) glycerol for several months without appreciable loss of activity.

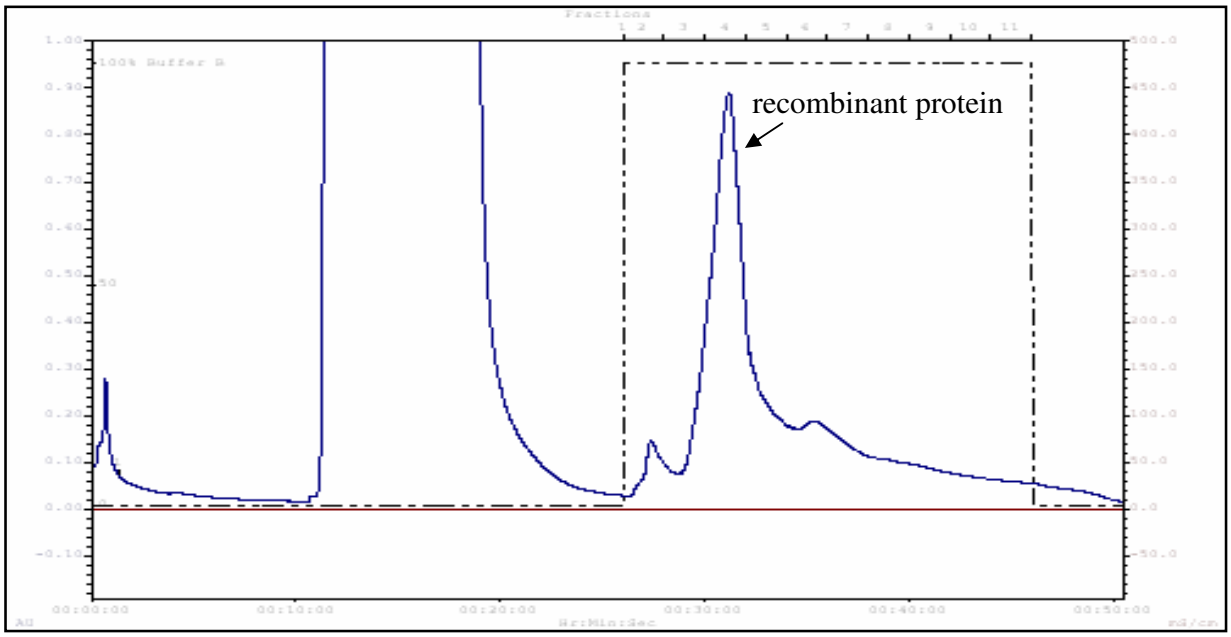


Fig. 4-11: Affinity chromatography of recombinant CoA ligase

4.8 Gel Electrophoresis (SDS-PAGE)

The grade of purity of the affinity-isolated protein was assessed by SDS–PAGE. Gel electrophoresis was carried out according to the method of Laemmli (1970), as described in 3.12.4. Crude protein extract before and after induction (10 µg each) and 2, 4, and 8 µg of the purified protein were analyzed in a 8% gel (Fig. 4-12). Protein bands were stained with Coomassie Brilliant Blue. The cell-free extract after IPTG induction exhibited a strong protein band at approximately 70 kDa (*lane* 2). After purification by affinity chromatography on the nickel column a single major band was visible, corresponding to a molecular mass of 70 kDa including the 6xHis tag. This molecular mass is somewhat higher than the value predicted from the amino acid sequence, i.e. 66,220 Da.

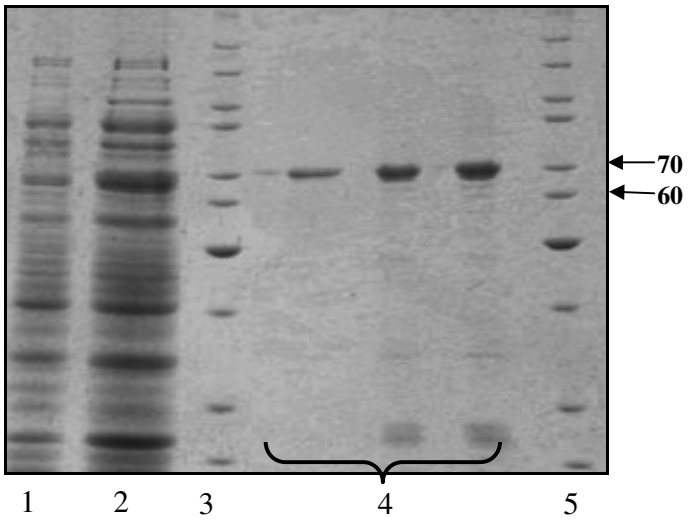


Fig. 4-12: SDS–PAGE of bacterial crude extracts and affinity-purified recombinant protein

lane 1 Soluble protein from uninduced BL21(DE3)pLysS cells (10 µg protein)
lane 2 Soluble extract from BL21(DE3)pLysS cells harboring the His-tag protein after ~18 h induction with IPTG (10 µg protein)
lane 3, 5 Molecular mass markers
lane 4 a,b,c Recombinant protein after purification on a His-trap column (2, 4, 8 µg protein).

4.9 Gel Filtration

The relative molecular mass of the native protein was determined by gel filtration on a Sephacryl S-200 High Resolution column using the buffer described in 3.14.7. The elution profile of the recombinant protein is shown in (Fig. 4-13). The column was first calibrated with the following standard proteins: albumin (66 kDa), alcohol dehydrogenase (150 kDa), β-Amylase (200 kDa), blue dextran (2000 kDa), carbonic anhydrase (29 kDa), cytochrome c (12.4 kDa). After applying the individual samples (2 ml each) to the column, a standard curve was calculated by plotting the molecular mass versus V_e/V_o for each standard protein. The relative molecular mass of the recombinant protein, as concluded from this standard curve (Fig. 4-14), was 70,000, indicating that the enzyme is active as a monomer.

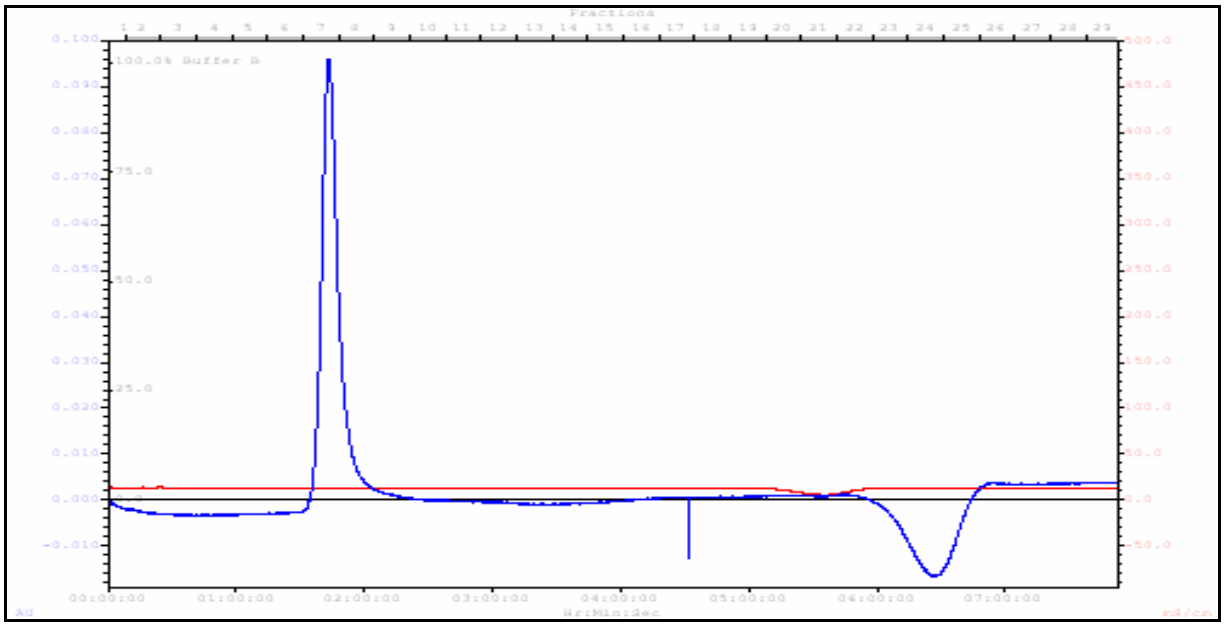


Fig. 4-13: Elution of the recombinant protein from a Sephacryl S-200 HR column. Detection was at 280 nm.

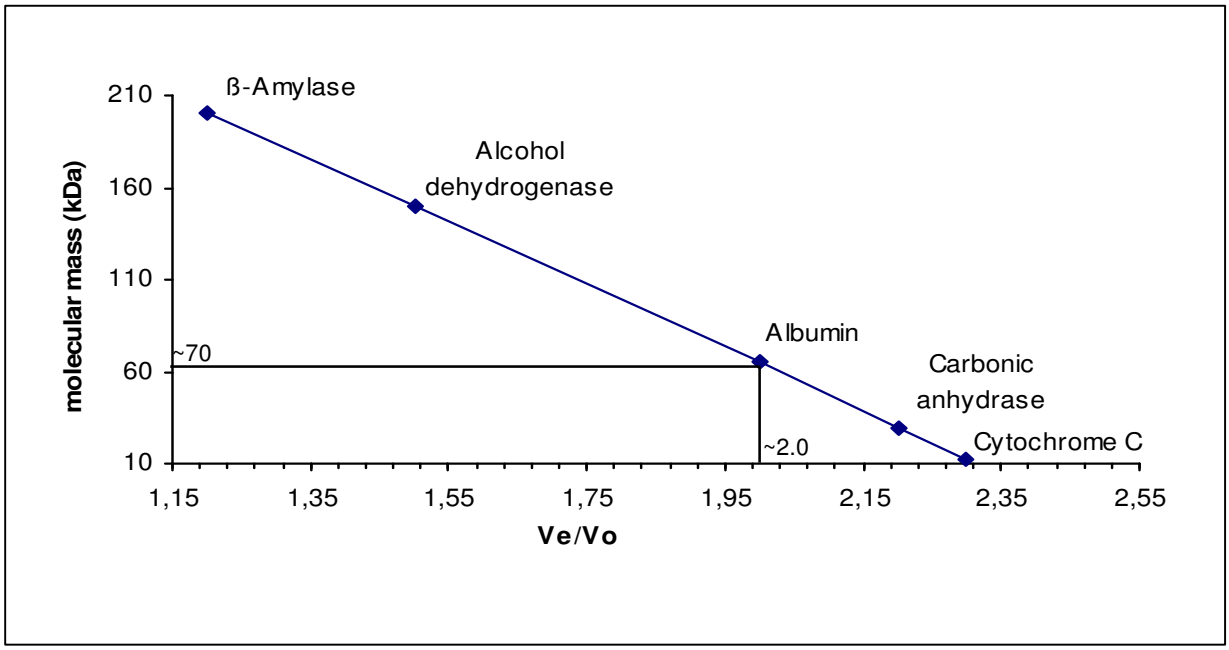


Fig. 4-14: Calibration of the Sephacryl S-200 HR column. The Ve/Vo for the recombinant CoA ligase was approx. 2.0.

4.10 Characterization of the Recombinant Enzyme

4.10.1 pH Optimum

The dependence of enzyme activity on pH is shown in Fig.4-15. 4CL activity was determined in 0.1 M potassium phosphate buffer covering the pH range from 4.5 to 8 in the standard assay. The optimum pH was around 6.5, which resembles that reported for other 4CLs. Half-maximal activity was observed at pH 5.7 and 7.5.

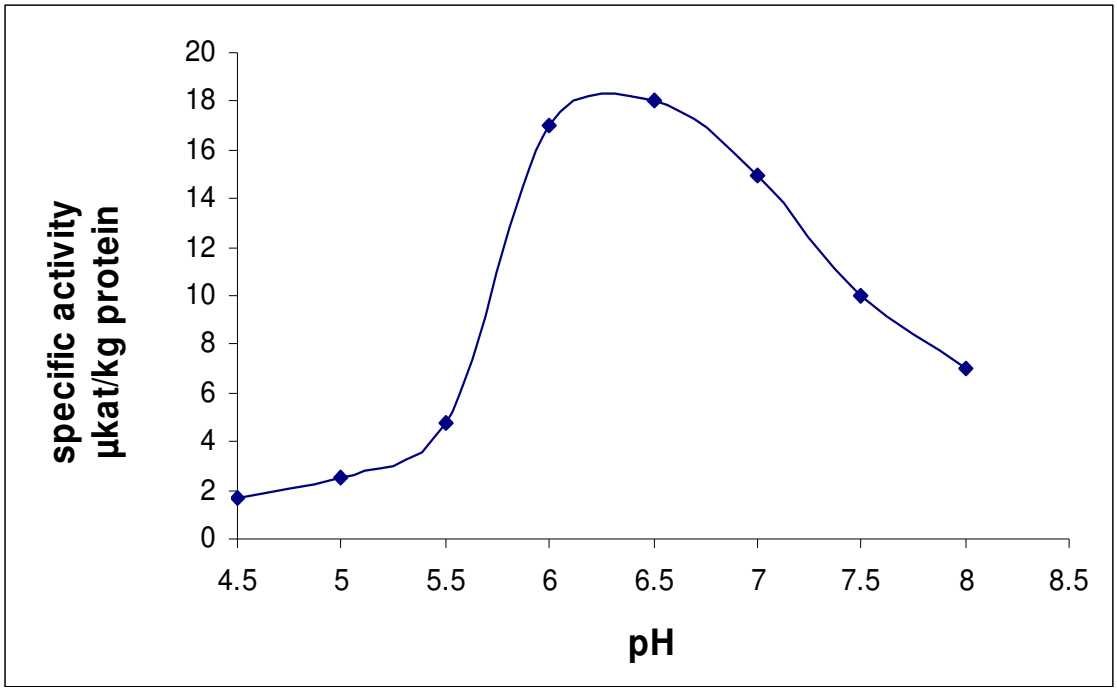


Fig. 4-15: pH optimum of the recombinant enzyme

4.10.2 Temperature Effect and Thermostability

To determine the temperature effect on the enzyme activity, standard assays were performed at 20, 25, 30, 35, 40, 45 and 50°C. The temperature optimum of the enzyme-catalyzed reaction was around 30°C. The activity dramatically decreased at temperatures higher than 35°C (Fig.4-16).

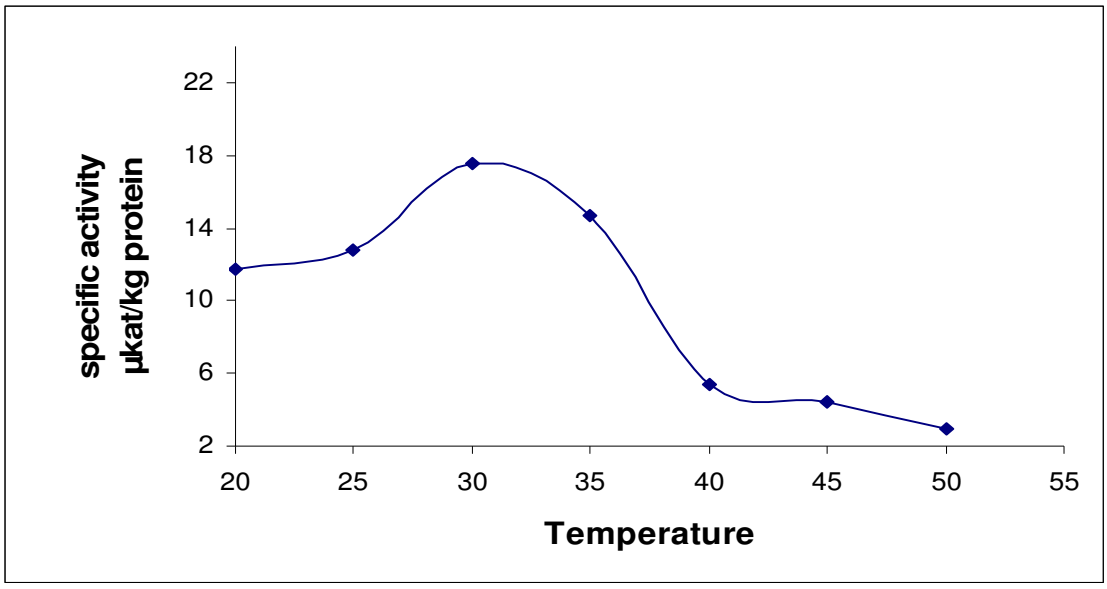


Fig. 4-16: Temperature optimum of the recombinant enzyme

4.10.3 Dependency of Enzyme Activity on Protein Concentration and Time

Plots of enzyme activity as a function of protein concentration and time are shown in Fig.4-17 and Fig.4-18. The reaction rate obtained with purified enzyme was approx. linear with the protein amount up to 8 µg in the standard assay and with time up to 15 min

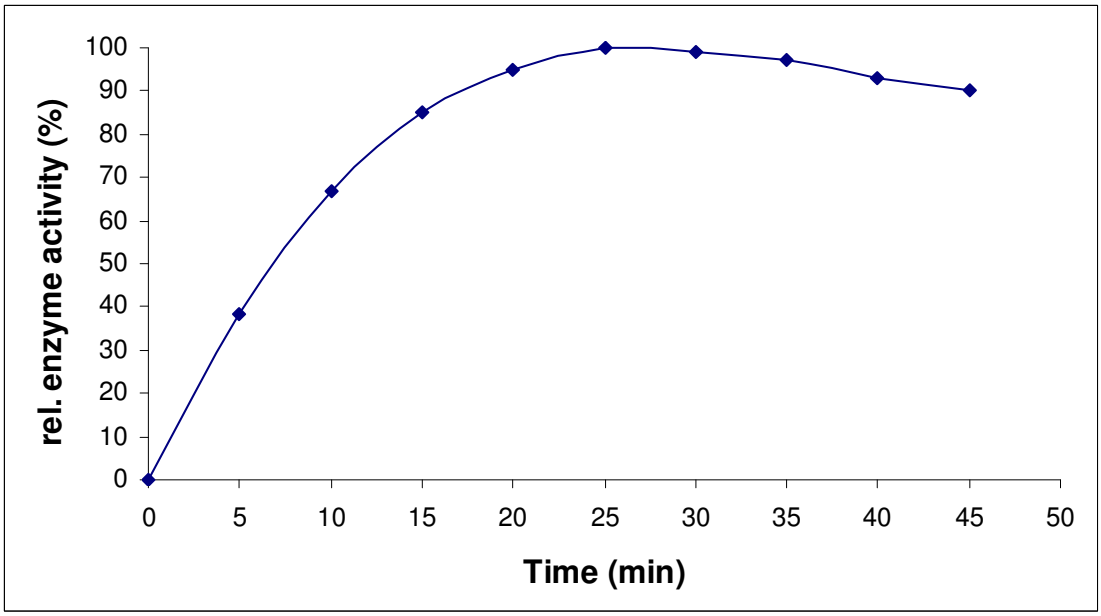


Fig. 4-17: Product formation as function of incubation time

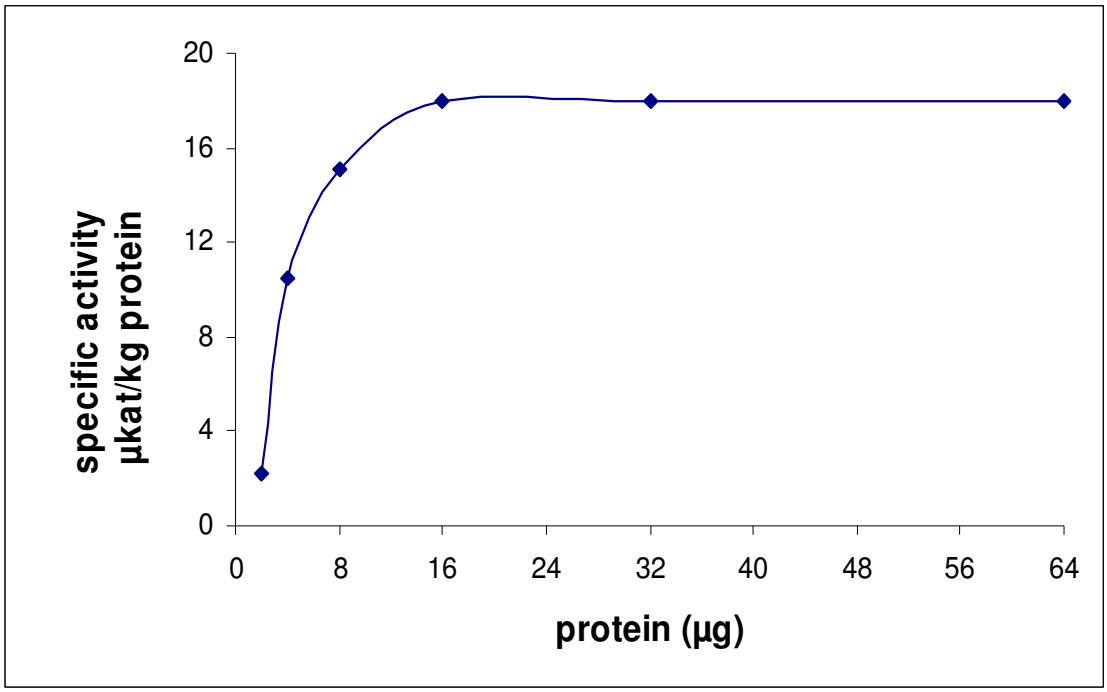


Fig. 4-18: Product formation as function of the protein amount in the standard assay

4.10.4 Substrate Specificity of 4-coumarate:CoA ligase

Under optimal pH, temperature, and protein concentration conditions, the FPLC-purified recombinant protein was tested for its ability to use different cinnamic and benzoic acids as substrates. Enzyme activities were first measured using a spectrophotometric assay at room temperature (3.1.15.1), which however is not applicable to benzoic acids. In order to determine the activity of the enzyme with benzoic acid derivatives, formation of the respective CoA esters in the reaction mixtures was examined by HPLC. For comparison, assays with cinnamic acids were also analyzed by HPLC. The recombinant protein showed similar patterns of substrate usage in the spectrophotometric assay and the HPLC analysis (Tab. 4-6 and Fig. 4-19). The relative substrate-utilization profiles were obtained using 0.2 mM substrate concentrations. Results are average values from three independent experiments and reported as percentage of the activity with 4-coumaric acid which is set as 100%. The enzyme efficiently converted cinnamate derivatives to the corresponding CoA esters, with a strong preference for 4-coumaric acid and

caffeic acid. Decreasing activities were found toward *m*-coumaric, *o*-coumaric, ferulic, and cinnamic acids. Activity with sinapic acid was not detectable. Thus, the recombinant protein is a 4CL. However, it also exhibited activity with benzoic acid and some substituted derivatives.

Tab. 4-6: Substrate specificity of recombinant *S. aucuparia* 4CL

Substrate	Rel. enzyme activity (%)
Cinnamic acid derivatives	
4-coumaric acid	100
caffeic acid	94
3-coumaric acid	85
2-coumaric acid	70
ferulic acid	55
cinnamic acid	35
sinapic acid	nc
Benzoic acid derivatives	
benzoic acid	12
3-hydroxybenzoic acid	7
2-hydroxybenzoic acid	5
4-hydroxybenzoic acid	nc
3,5-dihydroxybenzoic acid	nc
2-methoxybenzoic acid	nc
4-hydroxy-3-methoxybenzoic acid	nc
2,4-dichlorobenzoic acid	nc

nc = no conversion

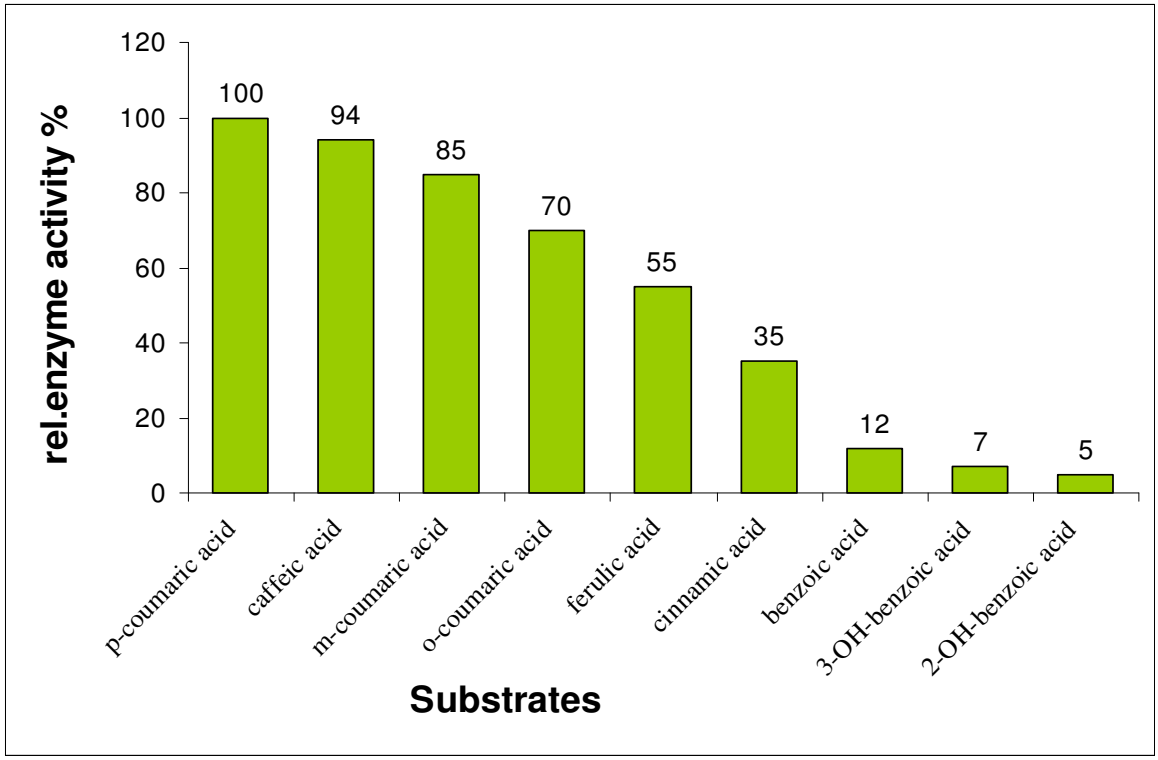


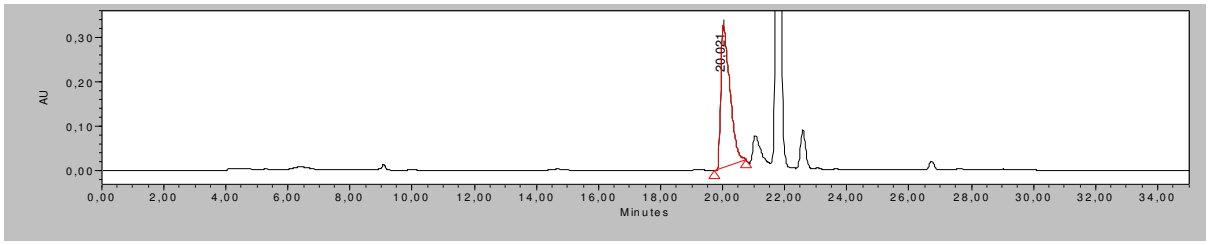
Fig. 4-19: Substrate utilization profile of recombinant *S. aucuparia* 4CL

4.10.5 Product Analysis by HPLC-DAD

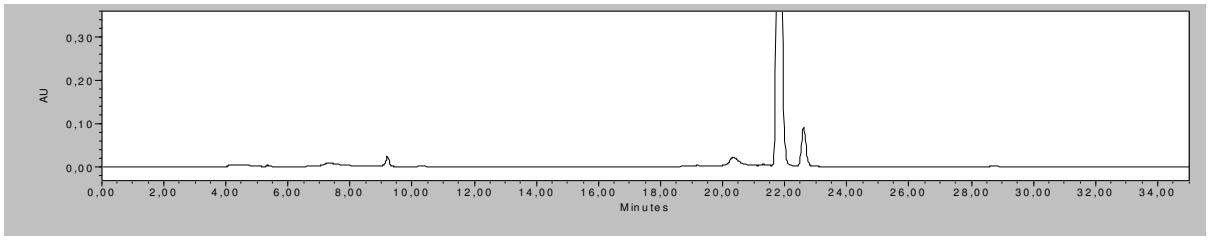
HPLC-DAD analysis was carried out as described in 3.15.2. The CoA esters formed by the purified enzyme from different substrates were identified by comparison of their Rt values and UV spectra with those of authentic compounds. Blanks for enzyme assays were performed with boiled enzyme. Chromatograms were recorded at 261 nm. Standard curves were constructed by plotting the peak areas of known quantities of the CoA ester reference compounds.

1. *p*-coumaric acid

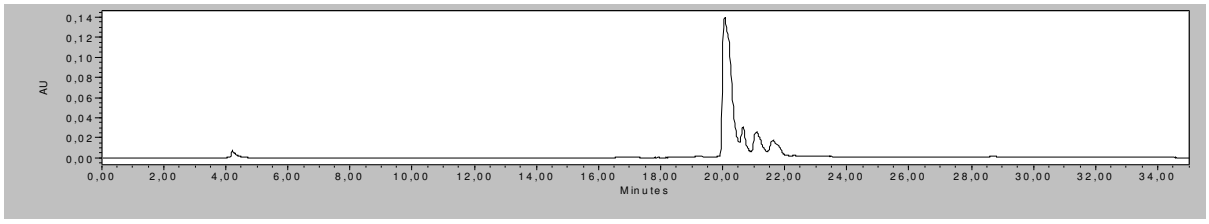
A)



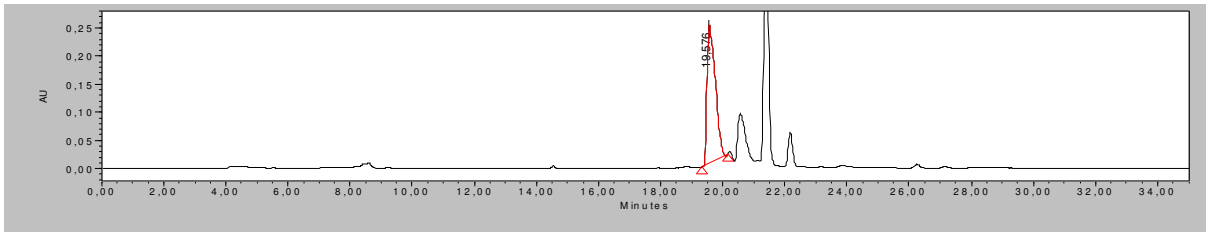
B)



C)



D)



E)

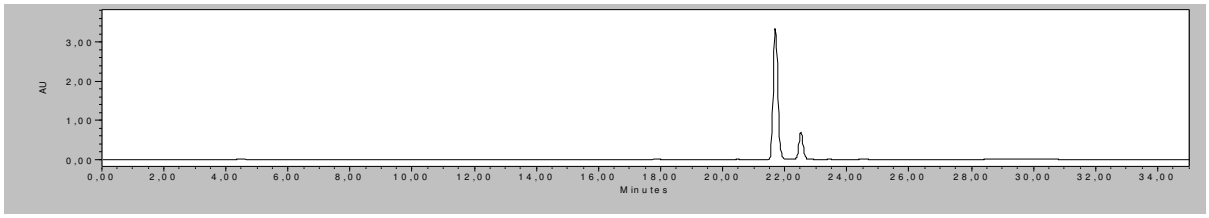
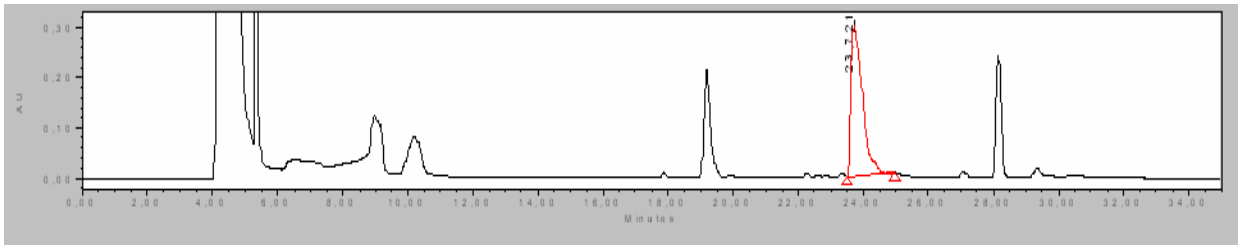


Fig. 4-20: HPLC analysis of enzyme assays with *p*-coumaric acid. Detection at 261 nm

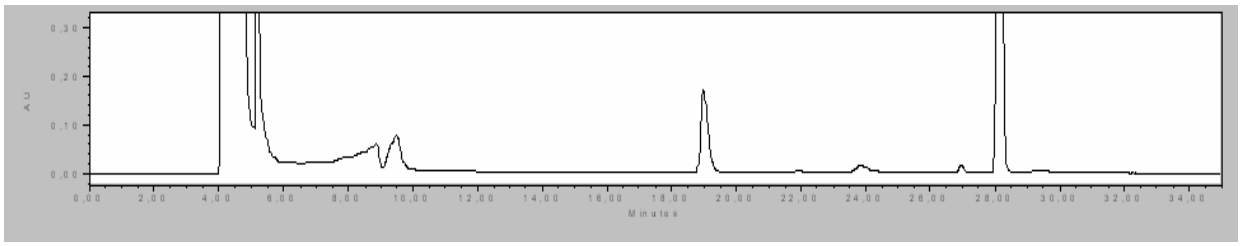
A) Incubation with active enzyme
B) Incubation with boiled enzyme
C) Reference compound *p*-coumaroyl-CoA
D) Co-chromatography (A + C)
E) *p*-coumaric acid (substrate)

2. Cinnamic acid

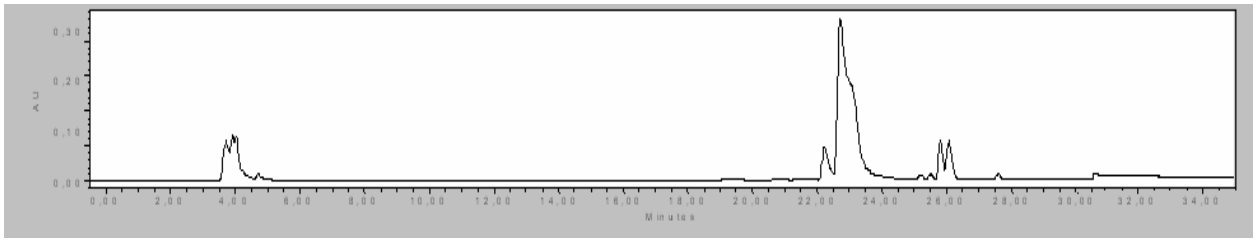
A)



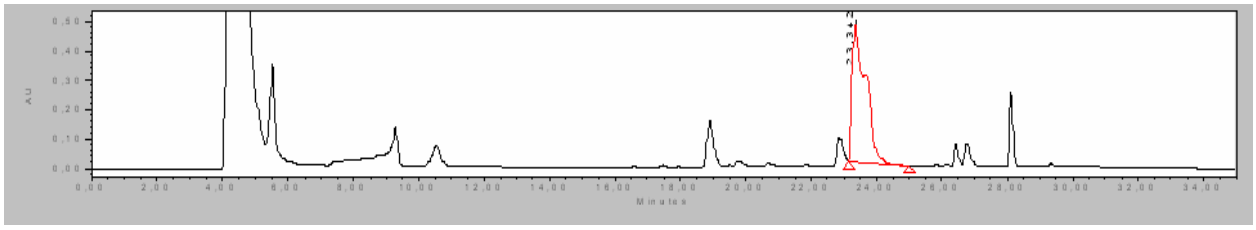
B)



C)



D)



E)

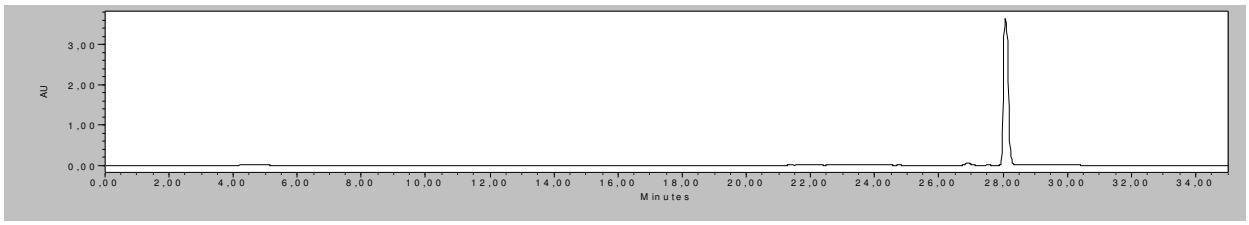
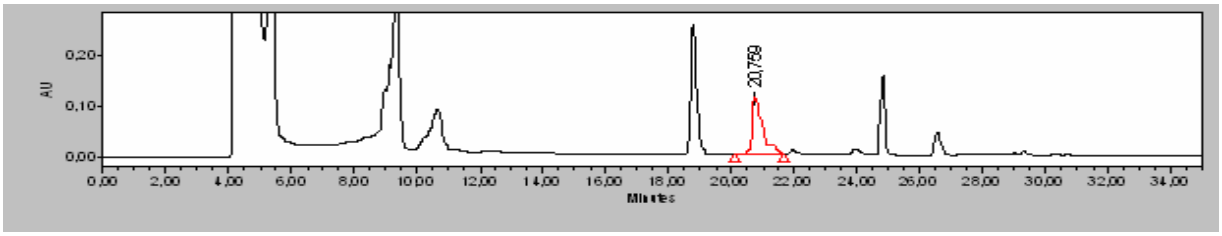


Fig. 4-21: HPLC analysis of enzyme assays with cinnamic acid. Detection at 261 nm

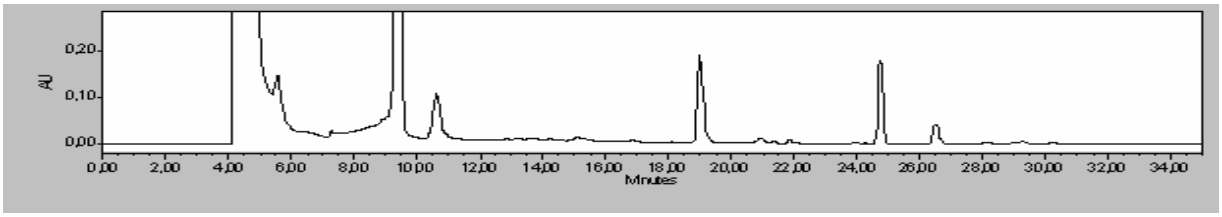
A) Incubation with active enzyme
B) Incubation with boiled enzyme
C) Reference compound cinnamoyl-CoA
D) Co-chromatography (A + C)
E) Cinnamic acid (substrate)

3. Benzoic acid

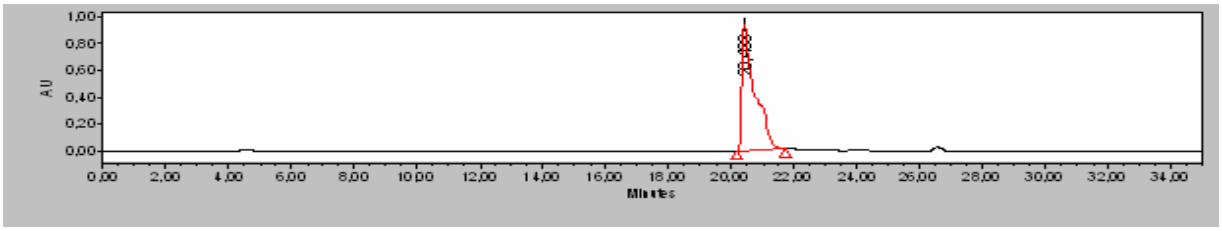
A)



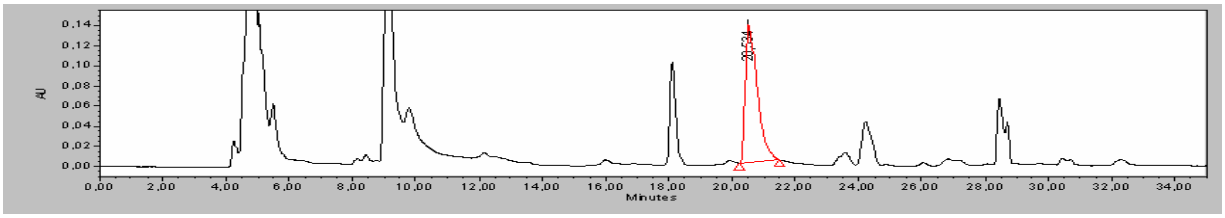
B)



C)



D)



E)

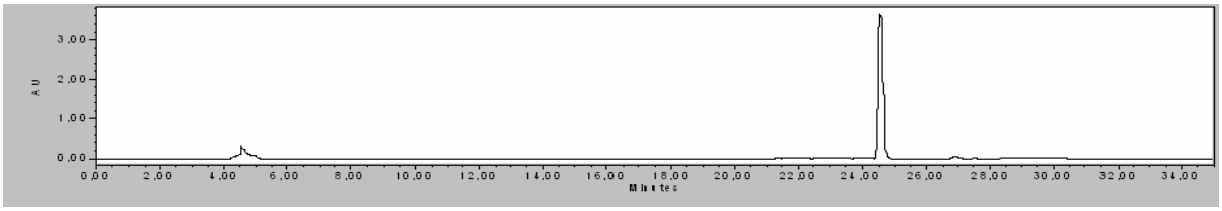
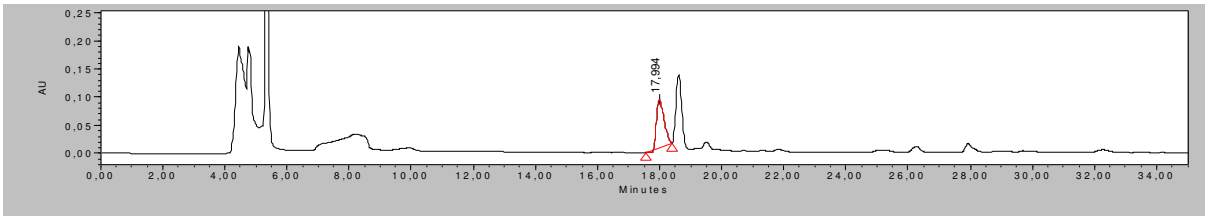


Fig. 4-22: HPLC analysis of enzyme assays with benzoic acid. Detection at 261 nm
A) Incubation with active enzyme B) Incubation with boiled enzyme
C) Reference compound benzoyl-CoA D) Co-chromatography (A + C)
E) Benzoic acid (substrate)

4. Caffeic acid

A)



B)

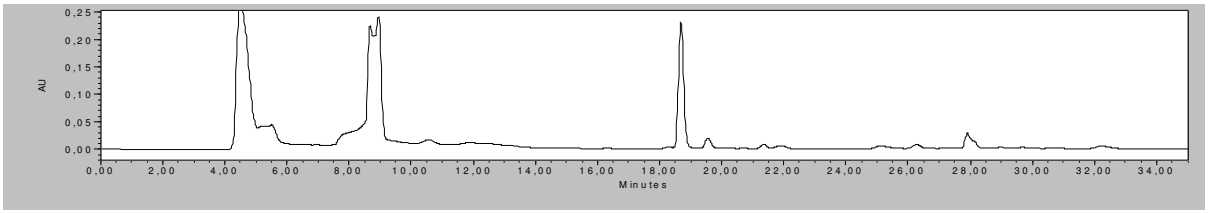
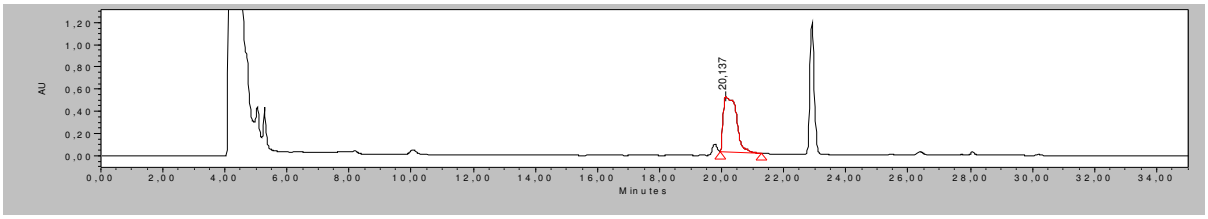


Fig. 4-23: HPLC analysis of enzyme assays with caffeic acid. Detection at 261 nm
A) Incubation with active enzyme B) Incubation with boiled enzyme

5. *m*-Coumaric acid

A)



B)

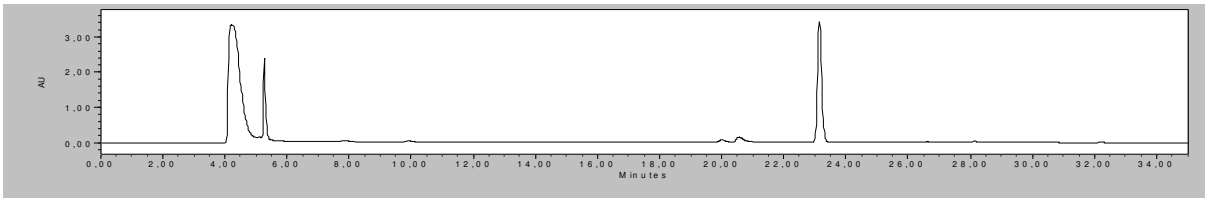
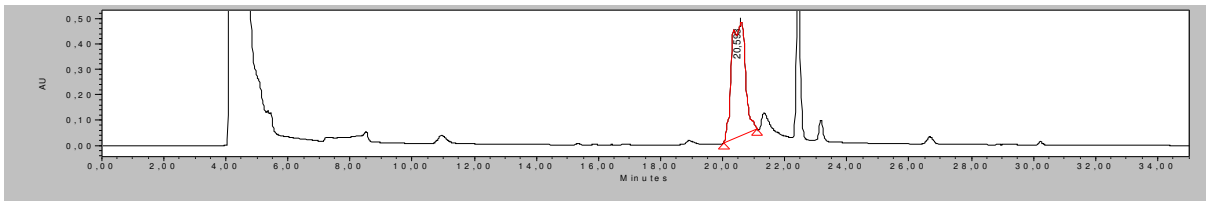


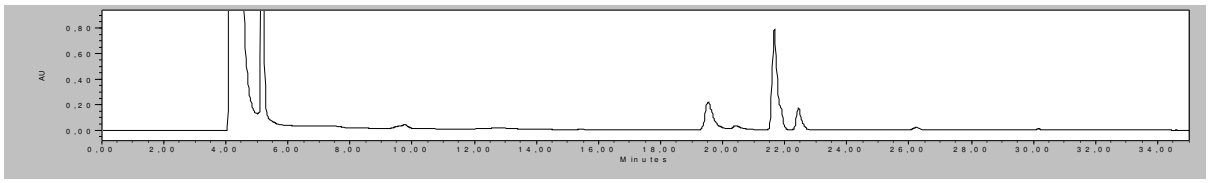
Fig. 4-24: HPLC analysis of enzyme assays with *m*-coumaric acid. Detection at 261 nm
A) Incubation with active enzyme B) Incubation with boiled enzyme

6. Ferulic acid

A)



B)



C)

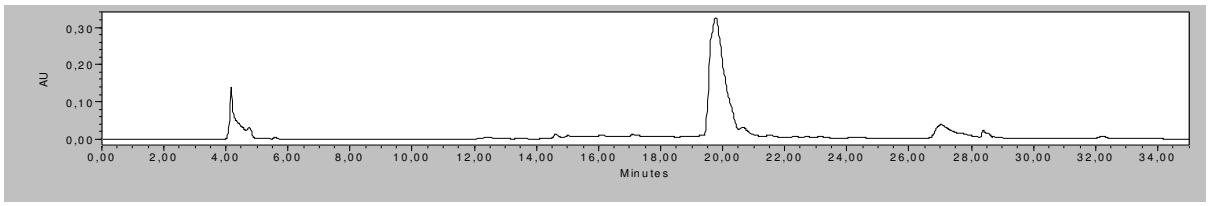
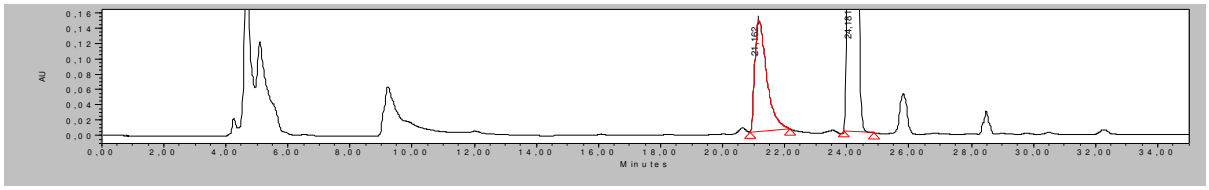


Fig. 4-25: HPLC analysis of enzyme assays with ferulic acid. Detection at 261 nm
A) Incubation with active enzyme B) Incubation with boiled enzyme C) Reference compound feruloyl-CoA

7. *o*-Coumaric acid

A)



B)

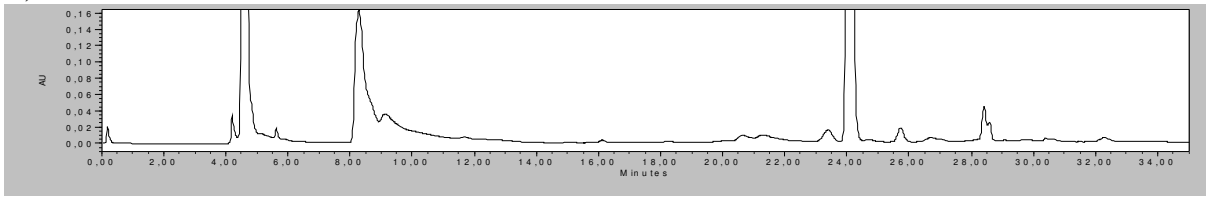
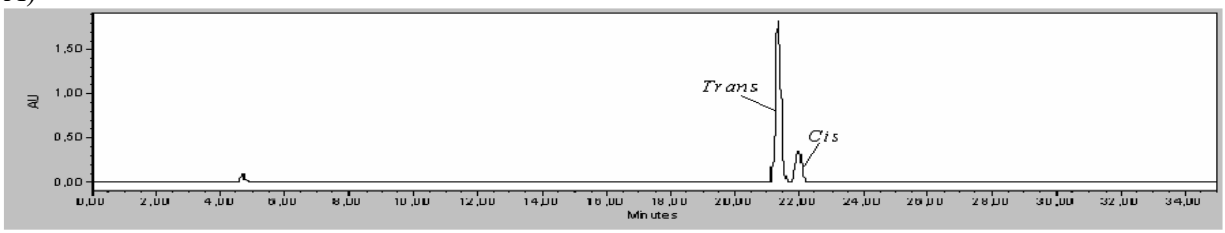


Fig. 4-26: HPLC analysis of enzyme assays with *o*-coumaric acid. Detection at 261 nm
A) Incubation with active enzyme B) Incubation with boiled enzyme

4.10.6 *Trans* to *Cis* Conversion of 4-coumaric Acid and 4-coumaroyl CoA

When incubations with 4-coumaric acid were analyzed by HPLC, this substrate and the resulting product, 4-coumaroyl-CoA, were usually observed as double peaks. Literature searches suggested these compounds to be the *trans* and *cis* isomers (Haskins et al. 1964; Ohashi et al. 1987; Braun and Tevini 1993; Locher et al. 1994; Rasmussen and Rudolph 1997; Wu et al. 2001; Yin et al. 2003). Their conversion was reported to be stimulated by UV light. Thus, incubations were carried out which were either exposed to or protected from UV light. With both 4-coumaric acid and 4-coumaroyl-CoA, *trans-cis* isomers were found. After irradiation of *trans*-4-coumaric acid with UV light, a strong increase in the amount of the *cis* isomer occurred, as shown in the HPLC profiles recorded before and after irradiation (Fig. 4-27).

A)



B)

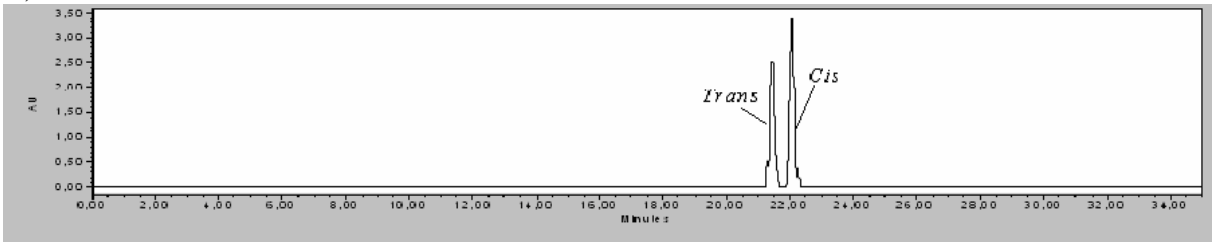
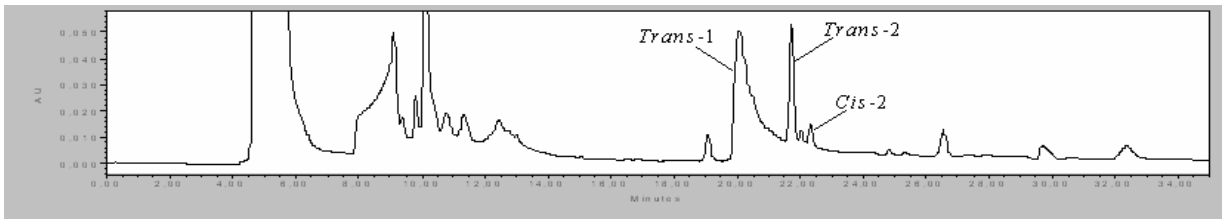


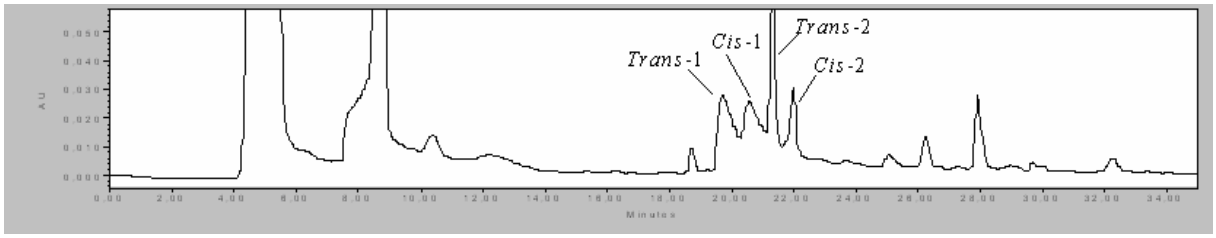
Fig. 4-27: *Trans–cis* conversion of 4-coumaric acid
A) 4-coumaric acid before UV light treatment B) 4-coumaric acid after a 2h-exposure to UV light at 366 nm

In order to study the effect of light on the *trans-cis* conversion of 4-coumaric acid and the corresponding CoA thioester, two independent enzyme assays with active 4CL were performed, with one tube being covered with aluminum foil to protect it from light. The control assay contained boiled enzyme. The samples were subjected to HPLC analysis, as described in (3.1.15.2). The following HPLC profiles show the *trans-cis* conversion for both 4-coumaric acid and 4-coumaroyl CoA (Fig. 4-28).

A)



B)



C)

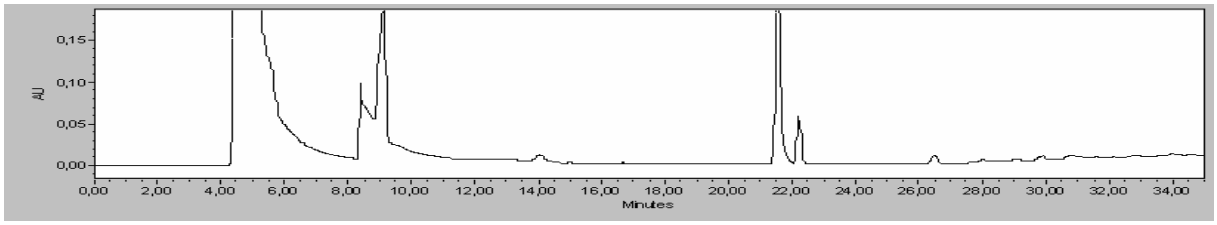


Fig. 4-28: HPLC analysis of *trans–cis* isomers in 4CL assays with 4-coumaric acid. Detection at 261 nm

A) 30 min-incubation wrapped in aluminum foil

B) 30 min-incubation without aluminum foil

C) 30 min-incubation with boiled 4CL, wrapped in aluminum foil

Trans-1= *trans*-4-coumaroyl-CoA *Trans*-2 = *trans*-4-coumaric acid

Cis-1 = *cis*-4-coumaroyl-CoA *Cis*-2 = *cis*-4-coumaric acid

4.10.7 Determination of Kinetic Data

Affinity-purified protein was used to determine the kinetic parameters. The increases in reaction rate with increasing concentrations of cinnamate and benzoate derivatives, ATP, and CoA were evaluated using spectrophotometry and HPLC. The Michaelis constant *Km* was estimated by linear regression of 1/*v* versus 1/[S], known as Lineweaver-Burk plot. The concentrations of the substrates were varied in the range from 5 µM to 640 µM for cinnamate derivatives, CoA and

ATP, whereas a range of 50 µM to 800 µM was chosen for benzoic acids. The kinetic parameters were calculated for product formation of the enzyme reaction at optimum pH, temperature, and enzyme amount. The experiments were carried out in triplicate using different concentrations of one substrate while the remaining reactants were saturating in the assay mixture. Fresh stock solutions of DTT, ATP, and CoA were made just prior to use, solutions of substrates were stored at -20 °C. Blanks for enzyme assays were performed with boiled enzyme. The following Lineweaver-Burk plots (Fig. 4-29 to 4-37) were generated to derive the apparent K_m values. For CoA, the Km value was 13 µM, for ATP it was 16 µM. The other data are listed in Tab. 4-7.

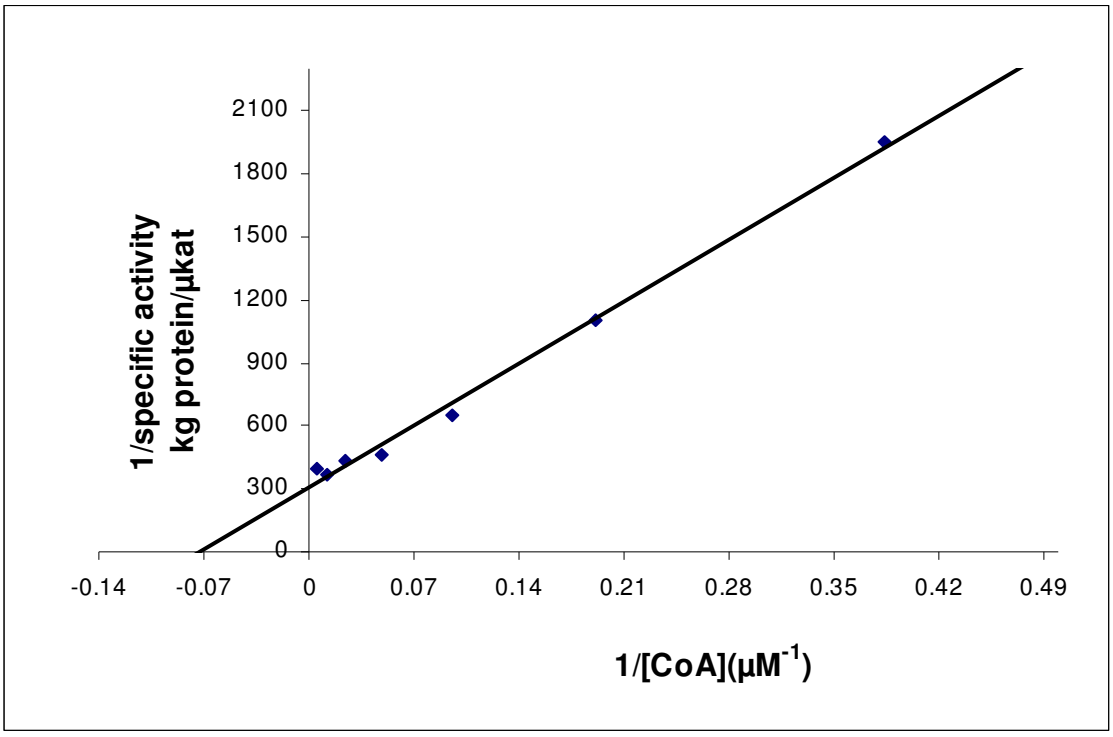


Fig. 4-29: Determination of the K_m value for CoA

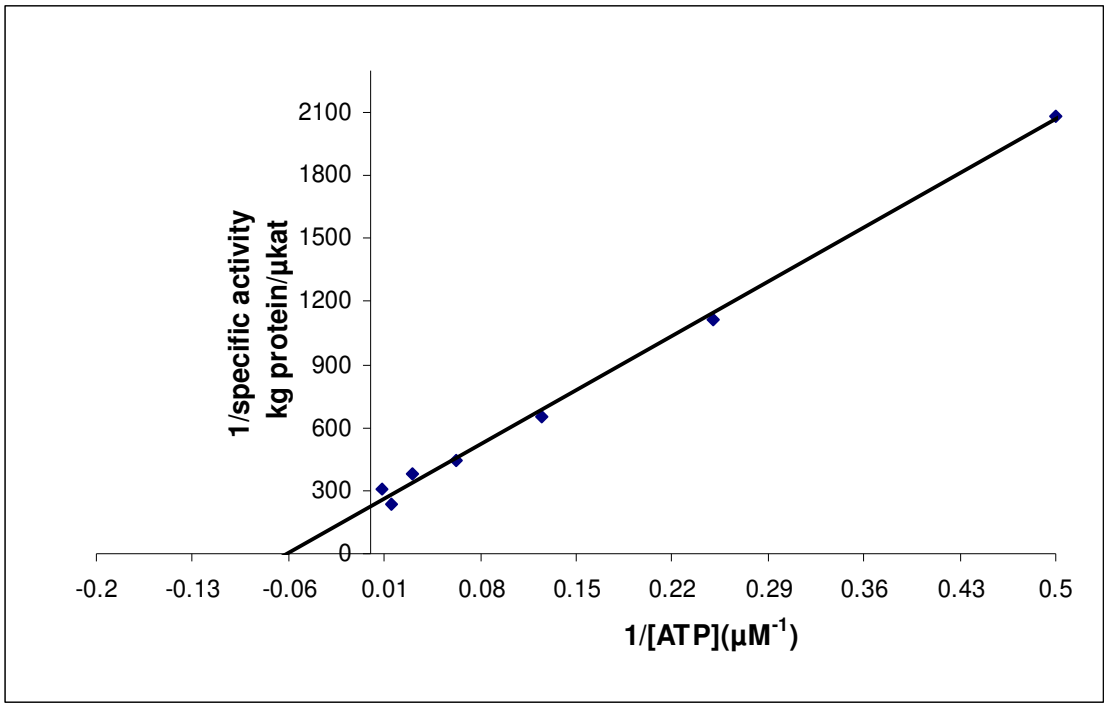


Fig. 4-30: Determination of the K_m value for ATP

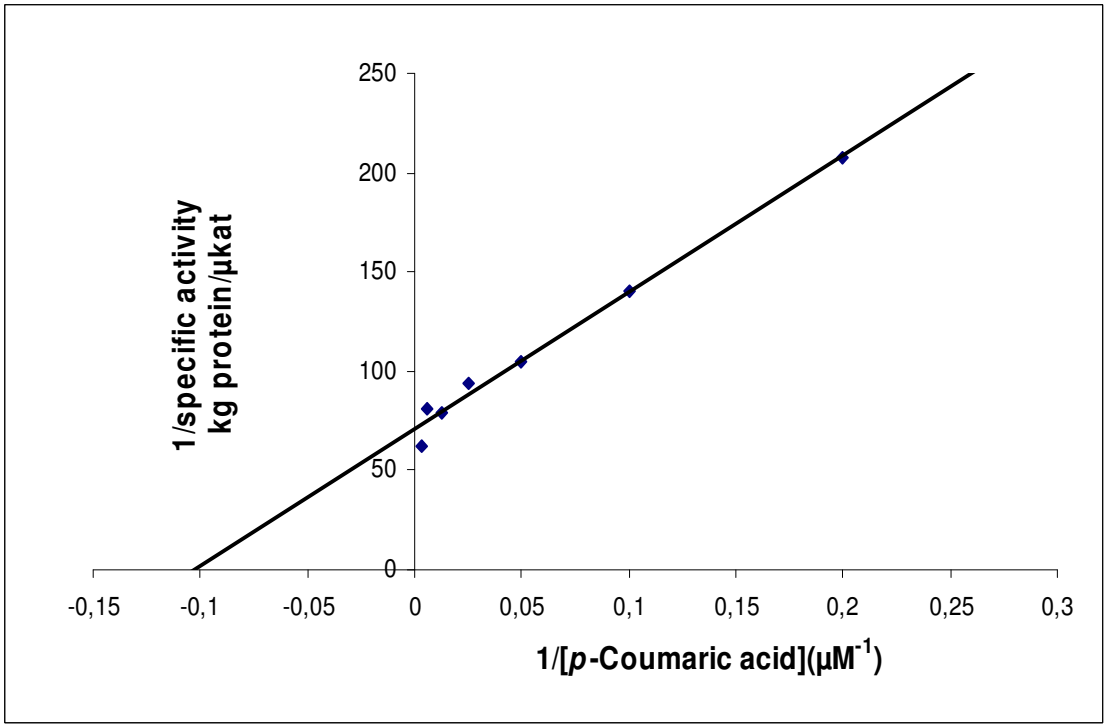


Fig. 4-31: Determination of the K_m value for *p*-coumaric acid

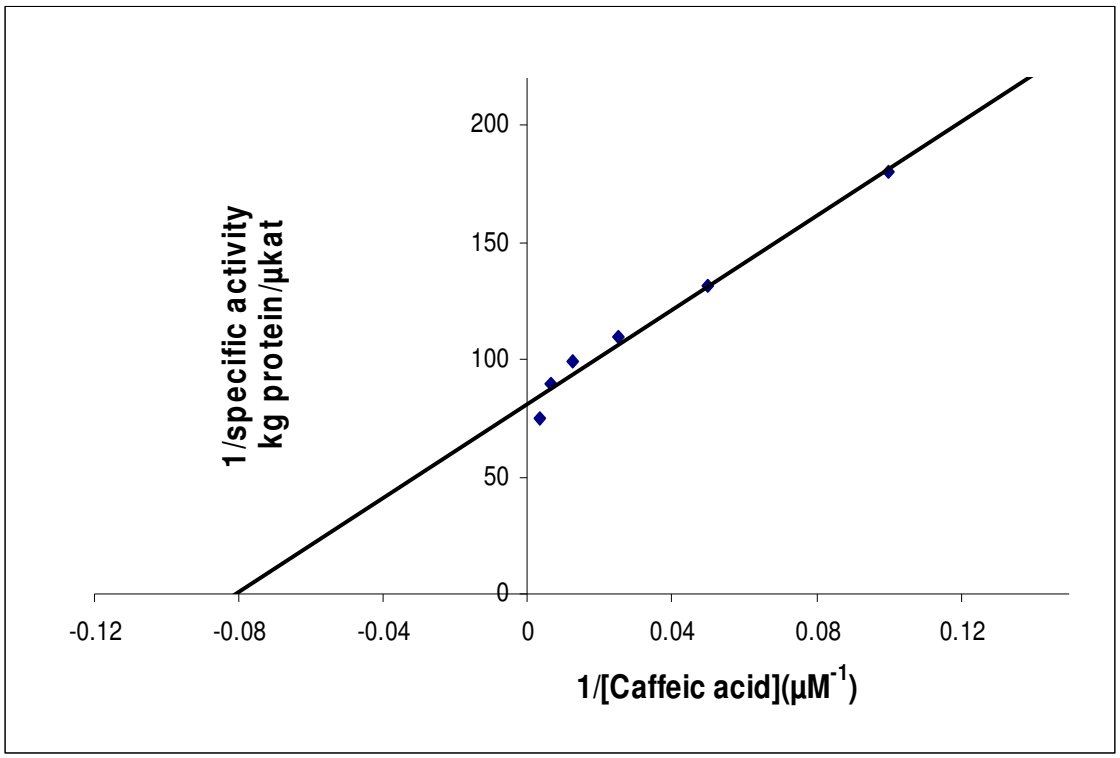


Fig. 4-32: Determination of the K_m value for caffeic acid

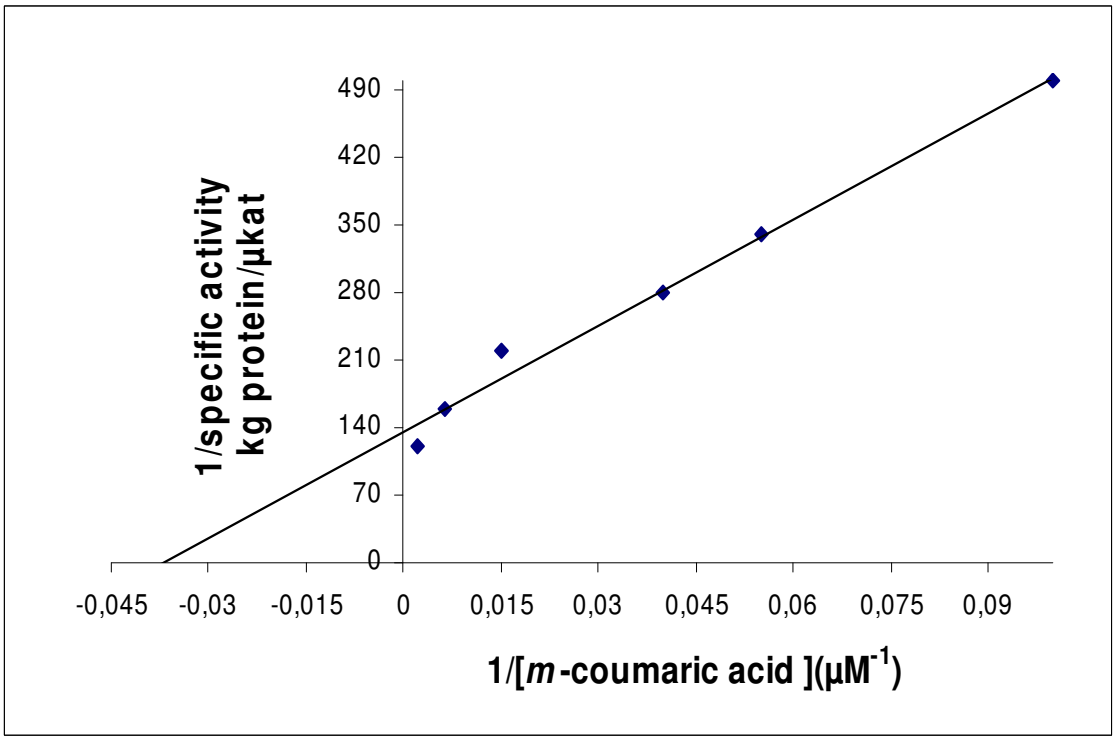


Fig. 4-33: Determination of the K_m value for *m*-coumaric acid

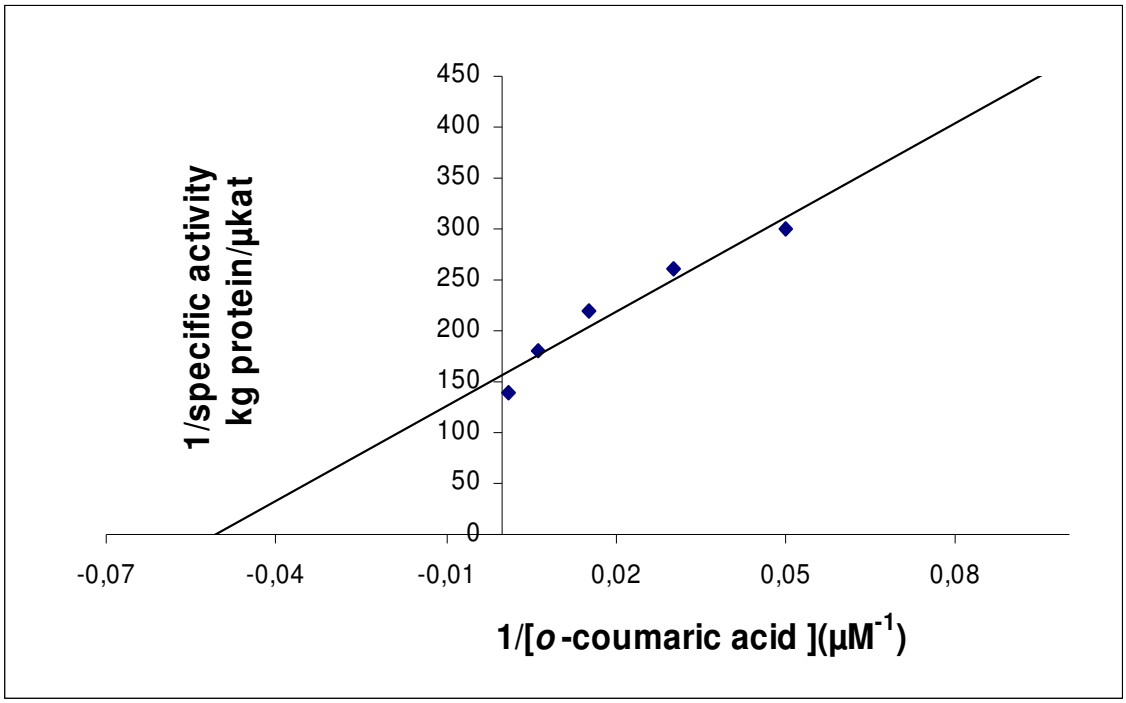


Fig. 4-34: Determination of the K_m value for *o*-coumaric acid

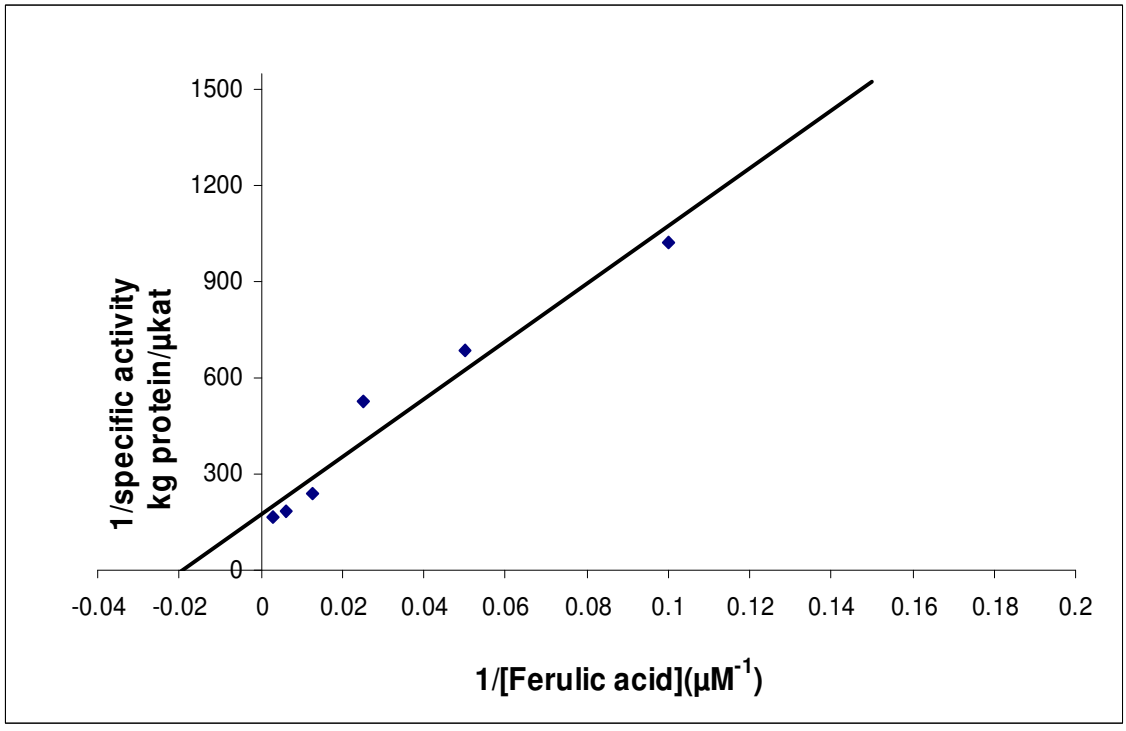


Fig. 4-35: Determination of the K_m value for ferulic acid

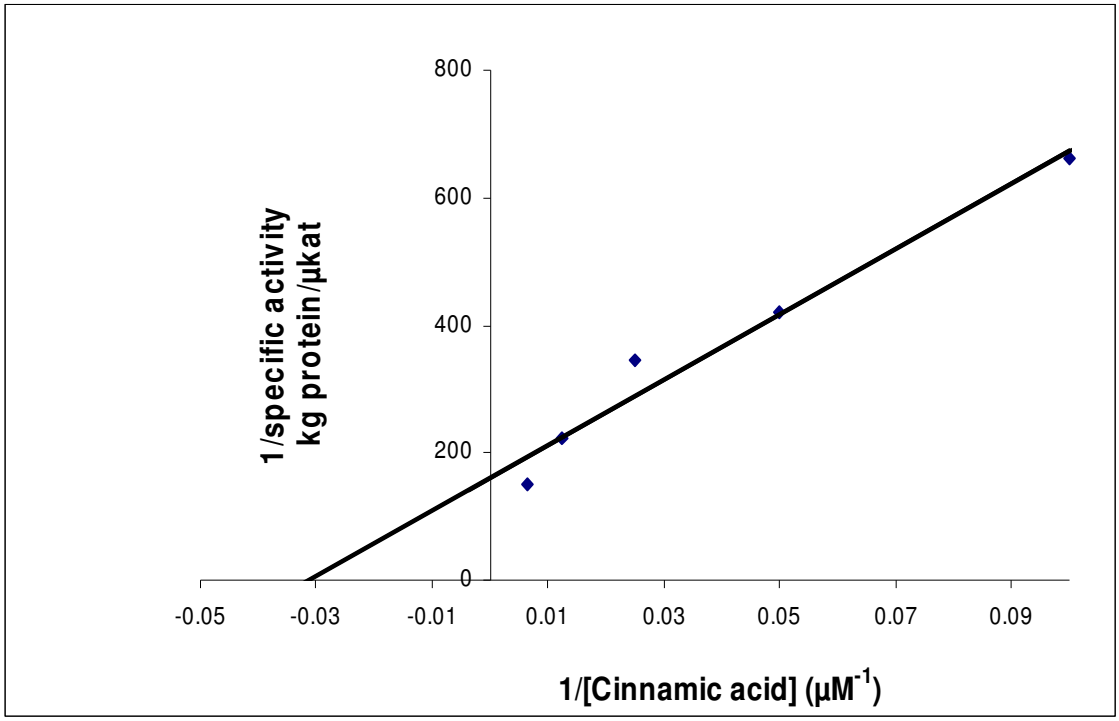


Fig. 4-36: Determination of the K_m value for cinnamic acid

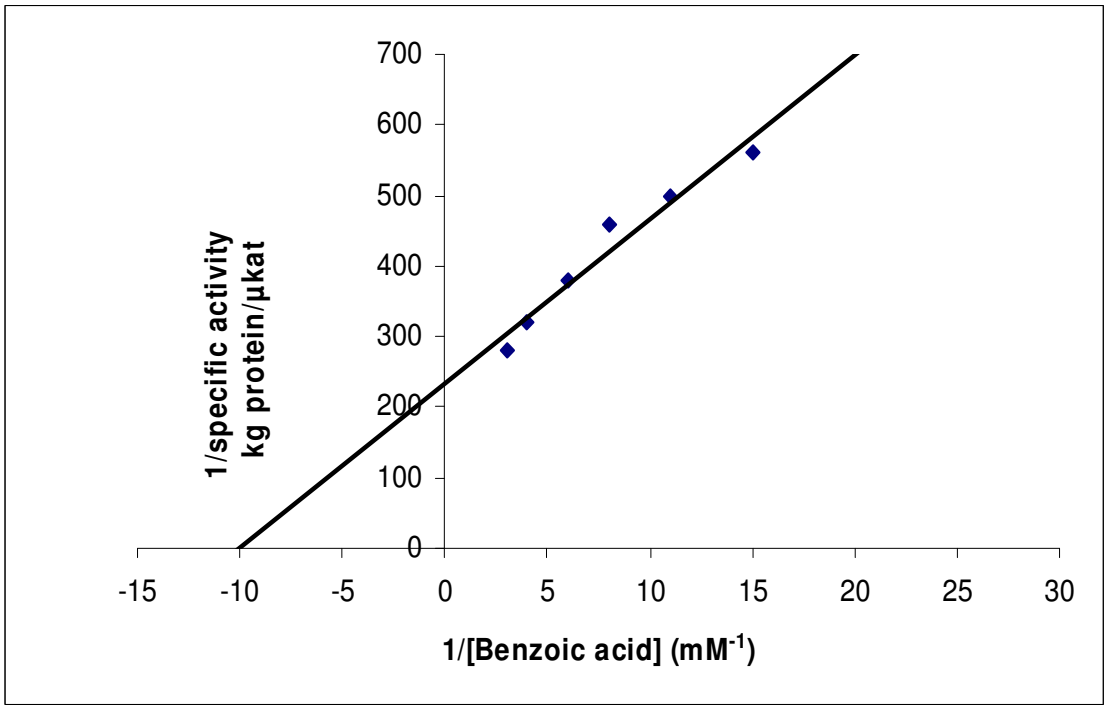


Fig. 4-37: Determination of the K_m value for benzoic acid

At saturating concentrations of all substrates and under the optimum conditions, V_{max}, K_{cat} and K_{cat}/K_m were determined on the basis of the Lineweaver-Burk plots. Table 4-7 lists these apparent kinetic parameters.

Tab. 4-7: Kinetic properties of 4CL from *Sorbus aucuparia* cell cultures

Substrate	K_m [µM]	V_{max} [pkat/µg]	K_{cat} [S^{-1}]	K_{cat}/K_m [M^{-1} S^{-1}]
p-coumaric acid	9 ± 3	16 ± 0.1	0.24	2.6×10^4
Caffeic acid	12.5 ± 2	12.30 ± 0.08	0.20	1.6×10^4
m-coumaric acid	28 ± 5	7.30 ± 0.1	0.12	7.5×10^3
o-coumaric acid	20 ± 3	6.40 ± 0.05	0.09	4.5×10^3
Ferulic acid	50 ± 7	5.70 ± 0.07	0.086	1.7×10^3
Cinnamic acid	33 ± 7	5.60 ± 0.08	0.080	2.4×10^3
Benzoic acid	110 ± 10	3.7 ± 0.2	0.057	5.2×10^2

The K_m values showed that the enzyme had the highest affinity for *p*-coumaric acid, followed by the other cinnamic acid derivatives. The highest K_m value was observed with benzoic acid. The k*cat* values demonstrated that 4CL achieved the highest rate of reaction with 4-coumaric acid and the lowest one with benzoic acid. That 4CL was catalytically most efficient with 4-coumaric acid, was confirmed by the K_{cat}/K_m values. The catalytic efficiency with benzoic acid was lowest.

5. DISCUSSION

Channelling of photosynthetically fixed carbon through the phenylpropanoid pathway in plants requires the sequential action of three enzymes, phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumarate:coenzyme A ligase (4CL), which together constitute the general phenylpropanoid pathway (Hahlbrock and Grisebach, 1979). As the last enzyme of this pathway, 4CL is responsible for the activation of cinnamic acid and its derivatives to their corresponding thioesters. These CoA esters are central intermediates in the biosynthesis of more highly modified phenylpropanoid compounds that are required for various physiological functions and for adaptation to environmental perturbations (Dixon and Paiva, 1995). 4CL gene expression is transcriptionally regulated and can be activated both during development and by external stimuli such as pathogen infection, elicitor treatment, methyl jasmonate treatment, wounding, and UV irradiation (Douglas et al., 1987; Schmelzer et al., 1989; Wu and Hahlbrock, 1992; Lee et al., 1995). Activation of 4CL by these stimuli was shown to occur co-ordinately with the activation of PAL, the first enzyme of general phenylpropanoid metabolism (Chappell and Hahlbrock, 1984; Logemann et al., 1995). 4CL is of central importance in the biosynthesis of many phenylpropanoid products, including lignin and flavonoids that play crucial roles in land plants. Thus, information on the evolution of 4CL genes will help to shed light on the evolution of phenylpropanoid metabolism associated with colonization of the land by plants. Furthermore, functional analysis of 4CL enzymes may help to reveal novel aspects of plant metabolism.

5.1 Reaction Mechanism Catalyzed by 4CL

4CL activates hydroxycinnamates by esterification with coenzyme A. The overall ATP- and Mg^{2+}-dependent reaction can be divided into the following two steps.

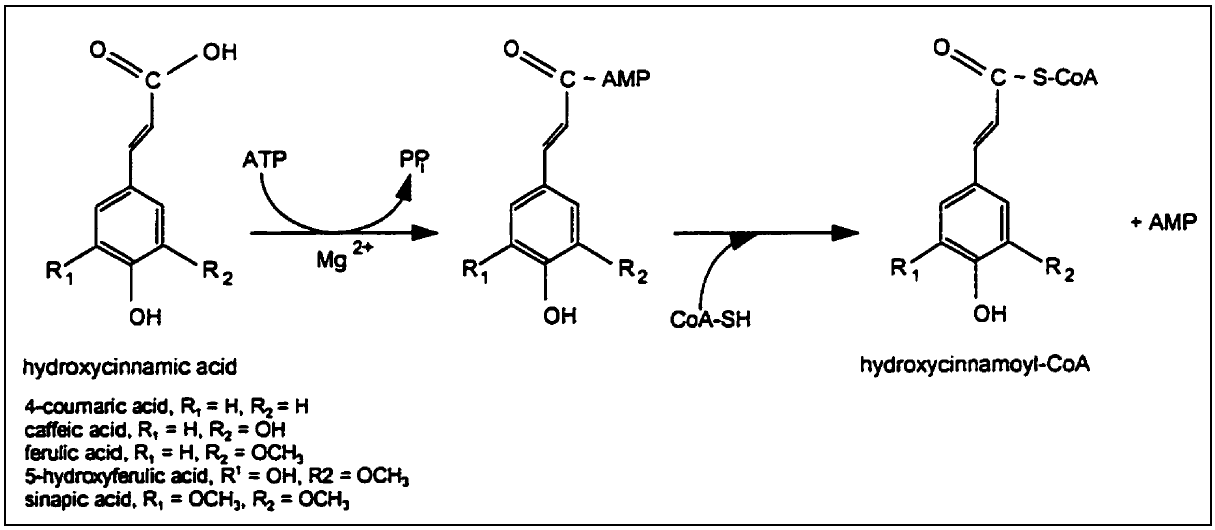


Fig. 5-1: Reaction catalyzed by 4-coumarate:CoA ligase

The first step is the formation of the adenylate intermediate. ATP apparently binds first to the enzyme followed by the substrate with pyrophosphate being released while the 4-hydroxycinnamoyl-adenylate intermediate still remains bound. Subsequent binding of coenzyme A (CoA) is followed by release of the 4-hydroxycinnamoyl-CoA thiol ester and AMP (Knobloch, and Hahlbrock., 1975; Becker et al., 1991). The carboxylate anion (a nucleophile) in 4-coumarate, caffeate, and ferulate is a non-activated, low energy resonance stabilized anion (the double bond is shared by both oxygens), making it extremely difficult to remove either of them enzymatically or non-enzymatically. The removal can be accomplished by activating the carboxylate group and introducing a good leaving group. The nucleophilic attack of the carboxylate anion to the electrophilic α-phosphorus of ATP generates 4-coumaroyl-AMP, a covalent adenylate intermediate, with concomitant release of pyrophosphate (Lehman, 1974; Engler and Richardson 1982). A pentacovalent phospharane adduct is apparently involved in the formation of the adenylate intermediate, which in turn serves as an activated intermediate for the subsequent reaction. Hydrolytic action of pyrophosphatases on pyrophosphate ensures functional unidirectionality under in vivo conditions, by removing one of the products of the first step (Randahl, 1979; Celis et al., 1985; Celis and Romero, 1987).

5.2 Relationship of 4CL to Other Adenylate-Forming Enzymes

Plant 4CLs, acyl-CoA ligases, peptide synthetases, and firefly luciferases are grouped in one family of AMP-binding proteins. Despite low overall amino acid sequence identity shared by these enzymes, similar reaction mechanisms and the presence of conserved peptide motifs were used as criteria to include them in a superfamily of adenylate-forming enzymes (Fulda et al., 1994). These enzymes do not only use a common reaction mechanism for the activation of carboxylate substrates but are also very likely marked by a similar functional architecture. The relationship of 4CL to other adenylate-forming enzymes was highlighted by functional analysis of key 4CL amino acids that are conserved in other adenylate-forming enzymes (Stuible et al., 2000). Within the superfamily, phylogenetic analyses show that 4CL forms a monophyletic plant-specific group that is most closely related to luciferases, when compared with long chain acyl-CoA synthetases and acetate-CoA synthetases (Cukovic et al., 2001). Two highly conserved peptide motifs within 4CL enzymes, box I as putative AMP-binding signature (PYSSGTTGLPKGV) and box II (GEICIRG), have been noted in comparisons of 4CL deduced amino acid sequences (Allina et al., 1998; Becker-André et al., 1991; Ehlting et al., 1999; Hu et al., 1998; Uhlmann and Ebel, 1993). The functional role played by box II, absolutely conserved in 4CLs, is unclear but is not directly involved in catalysis as previously assumed. Interestingly, the modification of highly conserved residues included in each of these boxes, which have been postulated to be essential for the enzymatic reaction, did not result in the total loss of activity but showed rather subtle changes of the kinetic parameters of the conversion of caffeate to the respective CoA ester (Stuible et al., 2000). Two mutations (Glu401Gln, Cys403Ala) introduced into the GEICIRG motif of At4CL2 (GenBank AAD47193) reduced the specific activity of the mutant enzymes to 21% and 45% of the wild-type levels, respectively, suggesting an important role for this highly conserved motif in enzymatic activity (Stuible et al., 2000). Both motifs, particularly box I, are also conserved among other members of the adenylate-forming enzyme superfamily. The regions of highest identity are functional domains, but significant identities can extend to other regions of the proteins (Cukovic et al., 2001).

The *S. aucuparia* 4CL sequence studied here contained near the N-terminus the conserved motif I (LPFSSGTTGLPKG) located between the amino acid positions 247 and 259 (Fig. 5-2). Towards the C-terminus, the second conserved sequence motif II (GEICIRG) was present

between the amino acid positions 447 to 453. Another conserved motif (GQGYGMTEA) is placed between the amino acid positions 391 and 399.

MISIASNSVETQKAADTATNLMPPLINSTSQQNLTKLQPPACTNNIIDSTTATATINHVFRSKLPDIA
IPNHLPLHTYCFQNLPEFSDRPCLIVGSTGKSYSFSETHLISQKTGAGLSNLGIQKGDVIMILLQNCA
EFVFAFMGASLIGAVTTTANPFYTTAEIFKQVKAANAKLIITQSQYVNKLREHPSSADGADQNNYPKL
GEDFKVVTIDDPPENCLHFSVLSEANEKELPDVVIDAEDPVA**LPFSSGTTGLPKGV**ILTHKNLVTSVA
QQVDGENPNLYLKEDDVVLCVLPLFHIFSFNSVLLCSLRAGAGVLLMHKFEIGTLLELIQRYRVSVAA
VVPPLVIALAKNPMVAEFDLSSIRVVLSGAAPLGKELEEALKSRVPQAVL**GQGYGMTEA**GPVLSMCM
AFAKEPMPTKSGSCGTVVRNAELKVLDPETGLSLGYNQP**GEICIRG**FQIMKGYLNDAAATATTIDTEG
WLHTGDVGYVDDDDEVFIVDRVKELIKFKGFQVPPAELESLLISHPSIADAAVVPQRDDAAGEVPVAF
VVRSDGLELTEEAVKEFIAKQVVFYKRLHKVHFVHAIPKSASGKILRKDLRAKLATATPPL

Fig. 5-2: Conserved motifs in the *S. aucuparia* 4CL amino acid sequence

5.3 4CL Gene Families and Substrate Specificities

In higher plants, 4CL is typically encoded by a gene family consisting of two to three members. In some species, such as *Arabidopsis* (Ehlting et al., 1999), aspen (Hu et al., 1998), and soybean (Uhlmann and Ebel, 1993), specific 4CL genes appear to be associated with the formation of unique phenylpropanoid end-products. In other species, multiple forms of 4CL appear to possess identical or nearly identical catalytic properties; examples are potato (Becker-André et al., 1991), parsley (Lozoya et al., 1988), pine (Zhang and Chiang, 1997), and hybrid poplar (Alina et al., 1998).

The range of substrates used by 4CL isoforms varies within and between plant species, but generally 4-coumarate (4-hydroxycinnamate) and caffeate (3,4-dihydroxycinnamate) are the best substrates, followed by ferulate (3-methoxy-4-hydroxycinnamate) and 5-hydroxyferulate (3-methoxy-4,5-dihydroxycinnamate), with cinnamate itself being a poor substrate (Douglas, 1996). 4CL activity towards sinapate (3,5-dimethoxy-4-hydroxycinnamate) has been described in a few cases in angiosperms (Knobloch and Hahlbrock, 1975; Kutsuki et al., 1981; Hamberger and Hahlbrock, 2004; Lindermayr et al., 2002). 4CL isoforms from parsley, tobacco and hybrid poplar have been expressed as recombinant proteins and all have similar substrate utilization profiles. These enzymes are capable of converting 4-coumarate, caffeate and ferulate, with very low activity towards cinnamate and no activity towards sinapate (Allina et al., 1998; Cukovic et al., 2001; Lee and Douglas, 1996; Lozoya et al., 1988). In

contrast, divergent aspen and *Arabidopsis* 4CL isoforms have been characterized that differ in their catalytic properties as well as their expression patterns. This has been taken as evidence that these isoforms have specialized functions in channelling phenylpropanoid metabolism to distinct branch pathways (Ehlting et al., 1999; Hu et al., 1998).

The *Arabidopsis* 4CL gene family consists of four members, At4CL1, At4CL2, At4CL3 and At4CL5 (Cukovic et al., 2001; Ehlting et al., 1999; Schneider et al., 2003; Hamberger and Hahlbrock, 2004). At4CL1 and At4CL2 are closely related to each other, while At4CL3 is more closely related to divergent 4CL genes from diverse angiosperms that are associated with flavonoid biosynthesis (Ehlting et al., 1999). At4CL1 and At4CL2 are unusual in that they encode 4CL isoforms with distinct catalytic properties (Ehlting et al., 1999). While At4CL1 utilizes both 4-coumarate and ferulate, At4CL2 is unable to convert ferulate and converts 4-coumarate with a much lower efficiency than At4CL1 (Ehlting et al., 1999). At4CL5 is apparently capable of ligating sinapic acid, as determined spectrophotometrically (Schneider et al., 2003). The same gene has again been reported as encoding a 4CL capable of ligating sinapic acid (Hamberger and Hahlbrock, 2004).

In soybean, the two discrete 4CL genes characterized, Gm4CL14 and Gm4CL16, have been proposed to correspond to two distinct 4CL isoforms identified in this species, consistent with the idea that discrete 4CL genes can encode function-specific isoforms (Knobloch and Hahlbrock, 1977; Uhlmann and Ebel, 1993).

In aspen, the Pt4CL1 and Pt4CL2 gene products are both structurally and functionally distinct, and the two genes are expressed in a compartmentalized manner (Hu et al., 1998). Pt4CL1 is expressed in lignifying xylem and the corresponding recombinant protein prefers substrates such as ferulic acid and 5-hydroxyferulic acid, suggesting that it plays a specific role in lignification (Hu et al., 1998). This proposed role for Pt4CL1 is corroborated by the results of antisense-mediated down-regulation of Pt4CL1 in aspen, which led to accumulation of substantially reduced levels of lignin (Hu et al., 1999). Pt4CL2, on the other hand, shows highest activity towards 4-coumaric acid, with little activity towards ferulic acid, and is expressed in leaf and stem epidermis but not in developing xylem tissue, suggesting that it may play a role in the biosynthesis of non-lignin phenylpropanoids (Hu et al., 1998).

Two 4CL cDNAs, predicted to encode proteins sharing 86% amino acid identity, were isolated from poplar clone H11, an F_1 hybrid from a *P. trichocarpa* and *P. deltoides* cross. Southern-blot analysis using DNA from the first two generations of a similar cross (Bradshaw et al., 1994) showed that these two cDNAs represent two different genes, 4CL1 and 4CL2, both of which are present in the parents of the hybrid clone. At least three, and possibly more,

forms of the enzyme were resolved using 4-coumarate as a substrate in preparations from the xylem of *P. euramericana* and *P. tremuloides* (Kutsuki et al., 1982; Meng and Campbell, 1997).

Based on kinetic analysis and the substrate utilization profile, *Sorbus* 4CL reported here resembles the enzymes described for many plants (Wallis and Rhodes, 1977; Lozoya et al., 1988; Voo et al., 1995; Lee and Douglas, 1996). It shows activity towards 4-coumaric acid, caffeic acid and ferulic acid, and has no activity against sinapic acid. Interestingly, a *trans-cis* conversion of 4-coumarate was observed in the *S. aucuparia* 4CL assays. *Cis*-cinnamic acids, like *cis*-4-coumaric acid, *cis*-ferulic acid, and *cis*-caffeic acid (Haskins et al. 1964; Ohashi et al. 1987; Braun and Tevini 1993; Locher et al. 1994; Rasmussen and Rudolph 1997; Wu et al. 2001; Yin et al. 2003), have been found as naturally occurring metabolites. The *trans*-to-*cis* conversion of cinnamic acids was observed both under light (Kahnt, 1967) and in dark-grown root tissues (Locher et al. 1994). Only few papers report the isolation and characterization of hydroxycinnamic acid residues present in a *cis* configuration, for example for Lamiaceae (Yoshida et al., 1990), Commelinaceae (Kondo et al., 1991), Liliaceae (Hosokawa et al., 1995a,b; Toki et al., 1998) and Solanaceae (Ando et al., 1999a,b; Slimestad et al., 1999). In vivo, the *trans* isomer always predominates. The change from *trans* to *cis* configuration of the terminal coumaric acid on malvidin produced some appreciable modifications both at the level of colour expression and colour stability (Florian et al 2001). *Trans*-to-*cis* isomerization of cell wall bound 4-coumaric acid was proposed to play a role in sensing of UV and blue light (Towers and Abeysekera, 1984).

Recombinant *Sorbus* 4CL readily converted cinnamic acid to the corresponding CoA ester with an efficiency of about 35% of that with 4-coumaric acid. In *Rubus idaeus*, Ri4CL1 exhibited relative activities of 100%, 27%, 22% and 11% towards 4-coumarate, ferulate, caffeate, and cinnamate, respectively (Amrita and Brian, 2003). In contrast, recombinant Ri4CL2 showed a strong preference for cinnamate (153%) and lower but still prominent activities towards 4-coumarate (100%), ferulate (96%) and caffeate (59%). Ri4CL3 was most active with 4-coumarate and displayed some activity towards caffeate (29%), but was inactive with both cinnamate and ferulate (Amrita and Brian, 2003). The capacity of naturally occurring 4CLs to activate cinnamic acid can be correlated with the hydrophobicity of their substrate binding pocket (SBP). In *Arabidopsis,* an increase in the hydrophobicity of the SBP resulted in At4CL2 variants with strongly enhanced conversion rates of cinnamic acid (Schneider et al., 2003).

Characterization of 4CL isoforms also indicated that they were able to convert *m*-coumarate and *o*-coumarate into the corresponding CoA esters (Douglas et al., 1987; Lozoya et al., 1988). Our data demonstrate that *S. aucuparia* 4CL activated *m*-coumaric and *o*-coumaric acids at high efficiencies of about 85% and 70%, respectively, compared to that with 4-coumaric acid.

In contrast, sinapic acid was not converted to sinapoyl-CoA by the recombinant *S. aucuparia* 4CL. As mentioned above, 4CL activity toward sinapate is quite low in many angiosperms (Lee et al., 1995, 1997; Lee and Douglas, 1996; Zhang and Chiang, 1997; Allina et al., 1998; Hu et al., 1998; Ehlting et al., 1999; Hamada et al., 2003). On the other hand, some 4CLs capable of converting sinapate have been reported in soybean (Knobloch et al. 1975; Lindermayr et al., 2002), poplar (*Poplus euramericana*) (Grand et al. 1983), and *A. thaliana* (Hamberger and Hahlbrock, 2004). Crude enzyme extract from the developing xylem of black locust clearly converted sinapate to sinapoyl CoA (Katsuyoshi et al., 2004). This is consistent with an early report that 4CL activity toward sinapate was detected in black locust using a spectrophotometric assay (Kutsuki et al., 1982).

Some studies have suggested alternative pathways for sinapyl alcohol biosynthesis in which 5-hydroxylation and subsequent *O*-methylation of the guaiacyl nucleus can occur through conifer aldehyde or alcohol in aspen and *Arabidopsis* (Matsui et al., 1994; Humphreys et al., 1999; Osakabe et al., 1999; Li et al., 2000). Aspen 4CL (Pt4CL1) selectively converted caffeate to its CoA ester in a reaction containing a mixture of cinnamate derivatives (Harding et al., 2002). This is also persuasive evidence that syringyl monolignol is synthesized through conifer aldehyde or coniferyl alcohol, but not through sinapate in plants. In a feeding study with shoots of black locust, deuterium-labeled sinapate was incorporated into syringyl lignin. This incorporation was also observed in oleander (*Nerium indicum*) but not in *Magnolia kobus* and *Arabidopsis thaliana*, suggesting that sinapate is a precursor of sinapyl alcohol in black locust and oleander and that syringyl lignin biosynthesis probably varies in angiosperms, depending on the species (Yamauchi et al., 2002; Yamauchi et al., 2003). In the case of soybean Gm4CL1, the absence of a single valine residue between Pro343 and Leu344 was confirmed to be important for the sinapate-converting activity by generating a mutant of *Gm*4CL2 and 3 with a deleted valine (Lindermayr et al., 2002). The absence of activity against sinapic acid, coupled with the apparent absence of catalytically distinct 4CL isoforms in poplar and other plants (Lozoya et al., 1988; Voo et al., 1995; Lee and Douglas, 1996), makes it unlikely that the differential expression of 4CL gene family members encoding enzymes with different substrate utilization profiles is a mechanism used in poplar, or most

other plants, to partition carbon into guaiacyl and syringyl lignin, or into other phenylpropanoid end products.

5.4 Formation of Benzoic Acids

The mechanism by which cinnamic acid is converted to benzoic acid has not yet been resolved at the molecular level. The 4CL enzymes so far characterized show no activity towards benzoic acids and no plant benzoate:CoA ligase cDNA has been identified. In contrast, benzoate:CoA ligases have been characterized, and their genes isolated, from benzoate-degrading microorganisms (Altenschmidt et al., 1991; Egland et al., 1995). In our laboratory, a CoA ligase catalyzing the esterification of benzoic acid and 3-hydroxybenzoic acid by coenzyme A has been detected in cell cultures of the medicinal plant *H. androsaemum* (Clusiaceae) (Schmidt and Beerhues 1997; Abd El-Mawla and Beerhues, 2002). Enzymatic activity that catalyzes the formation of 3-hydroxybenzoyl-CoA (an intermediate in the biosynthesis of xanthones) was found in crude extracts from cell cultures of *Centaurium erythraea* (Gentianaceae) and the protein responsible for this activity has been purified (Barillas and Beerhues, 1997; Barillas and Beerhues, 2000). This enzyme also exhibited some activity with benzoic acid (18% compared to that with 3-hydroxybenzoate). 3-Hydroxybenzoate:CoA ligase was monomeric, its molecular mass was estimated to be ~ 41 kDa. The *K*m values were 14.7 µM for 3-hydroxybenzoic acid, 8.5 µM for coenzyme A and 229 µM for ATP (Barillas and Beerhues, 1997). A similar enzyme was detected in bacteria and found to have 8% activity with benzoic acid compared to the activity with 3-hydroxybenzoic acid (Laempe et al., 2001). In *Clarkia breweri*, benzoate:CoA ligase activity was found to be present throughout the plant, it was purified from whole flowers and found to catalyze the thioesterification of benzoic acid to its corresponding CoA derivative in an AMP-forming manner, similar to 4CLs (Beuerle and Pichersky, 2002). The enzyme showed K_m values of 95, 130, and 45 µM for ATP, CoA, and benzoic acid, respectively. In addition to showing activity with benzoic acid, *C. breweri* benzoate:CoA ligase exhibited significant activity with other structurally related compounds, including anthranilic acid (50% relative activity) and 3-hydroxybenzoic acid (25% relative activity) (Beuerle and Pichersky, 2002).
Supporting evidence comes from studies examining 4-hydroxybenzoic acid biosynthesis from 4-coumarate in elicitor-treated carrot (*Daucus carota*) cell cultures and extracts (Schnitzler et al., 1992), potato tuber cell free extracts (French et al., 1976), and *Lithospermum.*

erythrorhizon cell extracts (Yazaki and Heide, 1991). CoA was required and 4-hydroxybenzaldehyde was detected. However, there has been no subsequent work to provide detailed confirmation of these findings. By contrast, in vitro studies with *L. erythrorhizon* supported the *β*-oxidation pathway from 4-coumaroyl-CoA (Löscher and Heide, 1994) as the major 4-hydroxybenzoic acid biosynthetic pathway. What might have caused this discrepancy in the results is largely unknown and care must be taken in interpreting these studies.

S. aucuparia 4CL was found to have activity toward benzoic acids. Benzoic acid, 2-hydroxybenzoic, and 3-hydroxybenzoic were activated at relative efficiencies of 12, 5, and 7%, respectively, compared to the activity with 4-coumaric acid. The K_m value for benzoic acid was 110 ± 10 µM and the Vmax value was 3.7 ± 0.2 pkat/µg protein. To our knowledge, this is the first plant 4CL exhibiting distinct activity with benzoic acids.

5.5 Phylogenetic Reconstruction

A previous phylogenetic reconstruction using 4CL amino acid sequences showed that 4CL gene family members in several species can be grouped into two major phylogenetically related clusters, class I and class II (Ehlting et al., 1999). This finding, based on relatively few species, suggests that an ancient 4CL gene duplication, prior to the divergence of the current angiosperm lineages, allowed the evolution of class I and class II gene family members. Based on the phylogenetic clustering, it was apparent that all dicot taxa have multiple class I and single class II-type 4CL genes. While class III consisted only of monocot 4CL sequences, this class was not supported by a high bootstrap value and, as such, might represent a subset of either class I or class II type sequences. Within the putative class III cluster, ryegrass (*Lolium perenne*) 4CL3 possesses a divergent motif II (GEICVRG), like class II *Rubus*4CL3. This suggests that even highly conserved domains within the 4CL sequences have undergone independent selective changes (Amrita and Brian, 2003). Within the class I cluster, there are two of the 4CL isoforms from *Arabidopsis* and genes from hybrid poplar, tobacco, soybean and potato.

We generated a new version of this phylogenetic reconstruction (Fig. 5-3) that incorporates the new amino acid sequence data of *S. aucuparia* 4CL. This analysis shows that *S. aucuparia* 4CL belongs to the class II angiosperm 4CL sequences and groups closely together with *Rubus* 4CL3 and soybean 4CL16. The distinctiveness of class II 4CL genes must reflect an ancient divergence from class I genes, since specific 4CL isoforms from highly diverged plant

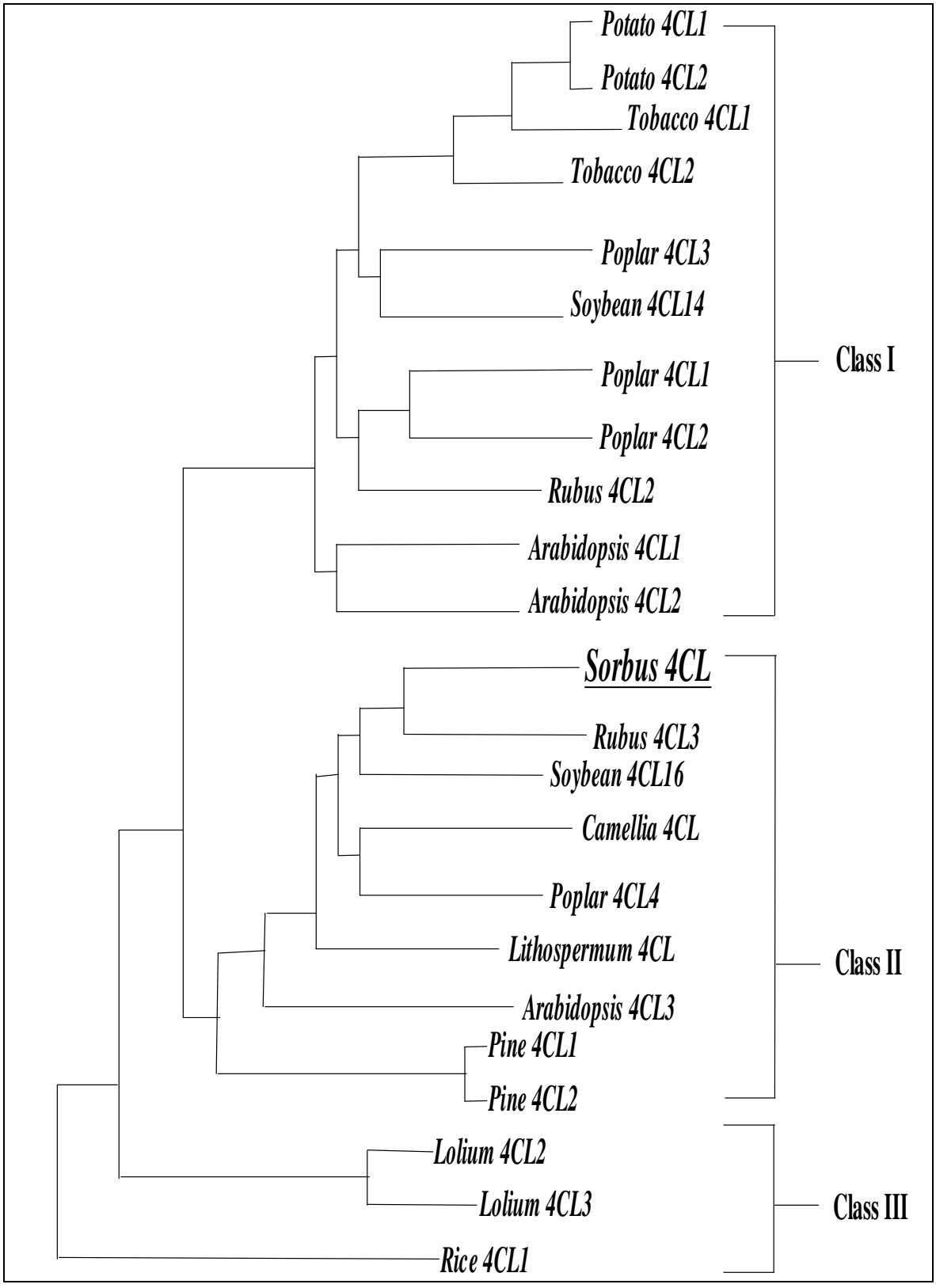


Fig. 5-3: Phylogenetic reconstruction based on 4CL amino acid sequences including *S. aucuparia* 4CL

taxa are found grouped together in class II rather than with other 4CL sequences from the same species. 4CL genes belonging to class II also share structural and functional differences from those in class I. For example, *Arabidopsis* 4CL3, *Lithospermum* 4CL, and now *Sorbus* 4CL, all in class II, have N-terminal extensions of 14–47 amino acids which are absent from class I sequences from the same species. It has been proposed that these variant N-terminal regions might be involved in phenolic substrate binding specificity (Hu et al., 1998). At4CL3 is more closely related to divergent 4CL genes from diverse angiosperms that are associated with flavonoid biosynthesis (Ehlting et al., 1999). The expression patterns of class I 4CL genes have also been shown to differ from those of class II 4CL genes in *Arabidopsis* (Ehlting et al., 1999) and soybean (Uhlmann and Ebel, 1995), suggesting that these genes might have acquired additional or unique functions that distinguish them from their progenitors. Important factors contributing to this lack of phylogenetic resolution may be the relatively narrow range of taxa for which data are available and the noncomprehensive isolation methods that have typically been used for detection of gene family members. From an evolutionary perspective, such gene duplication and divergence processes probably play an important role in the development of functional diversity by creating altered substrate specificities, as well as new interactions with effectors and/or activation processes.

5.6 Kinetic Properties and Biochemical Characterization

*K*m values reported for purified or partially purified 4CL proteins (Knobloch and Hahlbrock, 1975, 1977; Grand et al., 1983; Voo et al., 1995) range from 6.8 to 32 µM for 4-coumaric acid and from 9.1 to 130 µM for ferulic acid. Schneider et al. (2003) studied 4CL from *Arabidopsis* and determined Km = 233 ± 15 µM and Vmax = 475 ± 94 nkat/mg protein for 4-coumaric acid as well as Km = 22 ± 5 and Vmax = 236 ± 30 nkat/mg protein for caffeic acid. In the present study the values for 4-coumaric acid were Km = 9 ± 3 µM and Vmax = 16 ±0.1 pkat/µg protein, for caffeic acid Km = 12.5 ± 2 µM and Vmax = 12.34 ± 0.08 pkat/µg protein, for ferulic acid Km = 50 ± 7 µM and Vmax = 5.75 ± 0.07 pkat/µg protein. For cinnamic acid, the Km values obtained with 4CL isoforms from *Arabidopsis* range from 42 to 181 µM (Costa et al., 2005), whereas for the enzyme in this study the Km value was 33 ± 7 µM. Thus, the affinity of recombinant *Sorbus* 4CL toward cinnamic acid was higher than reported for any other 4CL enzyme and only 3.5-fold lower than the value for 4-coumaric acid. The affinities for cinnamic acid of 4CLs from other sources are 13-to 260-fold lower than those

for 4-coumaric acid (Knobloch and Hahlbrock, 1975; Knobloch and Hahlbrock 1977; Voo et al., 1995; Allina et al., 1998). The following table lists kinetic parameters for a number of 4CL enzymes (Table 5-1).

Tab. 5-1: Kinetic properties of 4CLs from various species including *S. aucuparia* 4CL

Species	Substrate	K_m (µM)	V_{max} (pkat/µg protein)
Arabidopsis (Costa et al., 2005)			
At4CL1	cinnamic acid	135 ± 4.9	9.2 ± 0.10
	p-coumaric acid	6 ± 0.4	61.3 ± 0.95
	caffeic acid	2 ± 0.1	42.0 ± 0.25
	ferulic acid	16 ± 1.3	39.5 ± 1.10
At4CL2	cinnamic acid	181 ± 4.3	3.2 ± 0.03
	p-coumaric acid	45 ± 2.4	46.6 ± 1.14
	caffeic acid	4 ± 0.4	30.4 ± 0.54
	ferulic acid	686 ± 32.8	9.9 ± 0.22
At4CL3	cinnamic acid	42 ± 1.8	4.0 ± 0.06
	p-coumaric acid	4 ± 0.2	12.9 ± 0.16
	caffeic acid	13 ± 1.0	7.0 ± 0.21
	ferulic acid	10 ± 0.7	6.4 ± 0.11
At4CL5	p-coumaric acid	25 ± 1.7	3.1 ± 0.07
	caffeic acid	7 ± 0.9	2.9 ± 0.09
	ferulic acid	6 ± 0.9	0.9 ± 0.03
	sinapic acid	10 ± 1.1	1.1 ± 0.04
Aspen (Harding et al., 2002)			
Pt4CL1	p-coumaric acid	55.64 ± 5.86	27.41 ± 1.97
	cinnamic acid	34.58 ± 4.63	17.49 ± 0.54
	ferulic acid	112.05 ± 6.73	15.80 ± 0.56
Pt4CL2	p-coumaric acid	3.57 ± 0.22	7.28 ± 0.31
	cinnamic acid	41.61 ± 1.20	6.80 ± 0.11
	ferulic acid	139.40 ± 6.37	4.92 ± 0.37
Poplar (Allina et al., 1998)			
4CL9	p-coumaric acid	80 ± 9	353 ± 36
	cinnamic acid	1048 ± 43	125 ± 3
	ferulic acid	102 ± 6	190 ± 11
Soybean (Lindermayr et al., 2002)			
Gm4CL1	cinnamic acid	4400	
	p-coumaric acid	22	
	caffeic acid	33	
	ferulic acid	8	
Gm4CL2	cinnamic acid	1700	

	p-coumaric acid	42	
	caffeic acid	13	
	ferulic acid	140	
Gm4CL3	cinnamic acid	1100	
	p-coumaric acid	9	
	caffeic acid	50	
	ferulic acid	3100	
Gm4CL4	cinnamic acid	260	
	p-coumaric acid	10	
	caffeic acid	34	
	ferulic acid	1300	
S. aucuparia (present study)			
Sa4CL	p-coumaric acid	9 ± 3	16 ± 0.1
	caffeic acid	12.5 ± 2	12.34 ± 0.08
	m-coumaric acid	28 ± 5	7.35 ± 0.1
	o-coumaric acid	20 ± 3	6.41 ± 0.05
	ferulic acid	50 ± 7	5.75 ± 0.07
	cinnamic acid	33 ± 7	5.60 ± 0.08
	benzoic acid	110 ± 10	3.7 ± 0.2

5.7 Prospective Studies

The work presented in this thesis is part of a molecular genetic project devoted to the elucidation of benzoic acid biosynthesis in elicitor-treated cell cultures of *Hypericum androsaemum* and *Sorbus aucuparia*. To date, no cDNA encoding a plant benzoate:CoA ligase has been isolated, despite the efforts by our own and other working groups. This is also true for most other biosynthetic enzymes involved. Identification of genes generally relies on the screening of cDNA/DNA libraries or the use of homology-based PCR approaches. Here, multiple cDNA fragments and a full-length clone were isolated. Continuation of this work will hopefully lead to the identification of two other full-length clones encoding cinnamate:CoA ligase and benzoate:CoA ligase. The substrate specificity-determining amino acid code of 4CLs was uncovered by detailed biochemical characterization of At4CL2, a naturally occurring loss-of-function mutant that is incapable of converting ferulic acid, and homology modeling of its substrate-binding pocket (Schneider et al., 2003). This specificity code allowed the rational design of At4CL2 variants with new catalytic properties, including

the capacity to activate ferulic, sinapic, and cinnamic acid, substrates not normally converted by the At4CL2 wild-type enzyme. Once all three CoA ligases from the above cell cultures have been cloned and their substrate-specificity-determining amino acids elucidated, structure-function-relationships will be studied by site-directed mutagenesis. This approach is also expected to lead to the generation of novel substrate preferences.

6. SUMMARY

► Innumerable plant-derived pharmacological agents contain benzoyl moieties. In addition, benzoic acids are among the regulators of the plant's interaction with its environment. Despite this importance and prevalence of plant benzoates, their biosynthesis is not well defined. The observed complexity is probably due to the simplicity of these molecules and their evolving chemical and functional elaboration. Several benzoic acid biosynthetic pathways are likely to exist in a given species, providing fine control over spatial and temporal synthesis and channeling intermediates to particular benzoate-derived products.

► A full-length cDNA sequence with homology to 4CL was identified from yeast-extract-treated *Sorbus aucuparia* cell cultures using degenerate primers derived from conserved regions of 4CLs and gene specific primers for RT-PCR. The nucleotide sequence of the composed clone and the properties of the recombinant enzyme were determined in this study. The sequence contained a 1815 bp open reading frame (ORF) encoding a polypeptide of 605 amino acids. The encoded protein was successfully expressed in *E. coli*. In addition, various cDNA fragments were cloned from both cell cultures and sequenced. They were identical in the overlapping portions when compared to the respective core cDNA fragments.

► FPLC on an affinity column was used to isolate the recombinant protein from bacterial crude protein extract. The N-terminal 6xHis-tag enabled purification on a His-trap column. The grade of purity of the affinity-isolated protein was assessed by SDS-PAGE. A single major band was visible, corresponding to a protein with a molecular mass of 70 kDa. On a calibrated gel filtration column, the relative molecular mass of the recombinant protein was ~ 70,000, indicating that the enzyme is active as a monomer. A standard curve had been calculated by plotting the molecular mass versus V_e/V_o for a series of standard proteins.

► The substrate preference of the affinity-purified recombinant *S. aucuparia* enzyme was tested. Most hydroxycinnamic acids were converted to the corresponding hydroxycinnamoyl-CoA thioesters, with 4-coumaric acid being the preferred substrate, followed by caffeic, *m*-

coumaric, *o*-coumaric, ferulic, and cinnamic acids. Thus, the recombinant enzyme is a 4CL. No evidence was found for the conversion of sinapic acid to sinapoyl-CoA.

► Interestingly, *S. aucuparia* 4CL was found to have activity toward benzoic acids. Benzoic acid, 2-hydroxybenzoic, and 3-hydroxybenzoic were activated at relative efficiencies of 12, 5, and 7%, respectively, compared to the activity with 4-coumaric acid. To our knowledge, this is the first plant 4CL showing distinct activity with benzoic acids.

► Kinetic analysis of recombinant *S. aucuparia* 4CL indicated that its relative abilities to use differently substituted hydroxycinnamic acids were similar to those observed with 4CLs from other species. For the preferred substrate 4-coumaric acid, the *Km* and *kcat/Km* values were 9 ± 3 and 2.6×10^4, respectively. For CoA and ATP involved in the reaction, *Km* values of 16 µM and 13 µM, respectively, were determined.

► Based on the amino acid sequences of 4CLs, a phylogenetic tree was constructed. *S. aucparia* 4CL belongs to the class II angiosperm 4CL sequences and groups closely together with *Rubus idaeus* 4CL3 and soybean 4CL16.

7. REFERENCES

Aas, G., Holdenrieder, O., 1992. Die Vogelbeere (*Sorbus aucuparia* L.) – der vielseitigePionier.Wald Holz 15, 48–50

Abd El-Mawla, Beerhues, L., 2002. Benzoic acid biosynthesis in cell cultures of *Hypericum androsaemum*. Planta, 214, 727-733.

Abd El-Mawla, Schmidt, W., Beerhues, L., 2001. Cinnamic acid is a precursor of benzoic acids in cell cultures of *Hypericum androsaemum* L. but not in cell cultures of *Centaurium erythraea* RAFN. Planta, 212, 288-293.

Akada, S., Dube, SK., 1995. Organization of soybean chalcone synthase gene clusters and characterization of a new member of the family. Plant Mol Biol 29(2); 189-99. PMID: 7579172

Allina, S.A., Pri-Hadash, A., Theilmann, D.A., Ellis, B.E. and Douglas, C.J.,1998. 4-Coumarate: conezyme A ligase (4CL) in hybrid poplar (Populus trichocarpa X Populus deltoides. Properties of native enzymes, cDNA cloning and analysis of recombinant enzymes. Plant Physiol. 116, 743-754.

Altenschmidt, U., Oswald, B., Fuchs, G., 1991. Purification and characterization of benzoate-coenzyme A and 2-aminobenzoate-coenzyme A ligases from a denitrifying *Pseudomonas* sp. J. Bacteriol. 173,5494-5501.

Ando, T., Saito, N., Tatsuzawa, F., Kakefuda, T., Yamakage, K., Ohtani, E., Koshi-ishi, M., Matsusake, Y., Kokubun, H., Watanabe, H., Tsukamoto, T., Ueda, Y., Hashimoto, G., Marchesi, E., Asakura, K., Hara, R., Seki, H., 1999a. Floral anthocyanins in wild taxa of Petunia (Solanaceae). Biochem. Systematics Ecol. 27, 623– 650.

Ando, T., Saito, N., Tatsuzawa, F., Kakefuda, T., Yamakage, K., Ohtani, E., Koshi-ishi, M., Matsusake, Y., Kokubun, H., Watanabe, 794 F. George et al. / Phytochemistry 57 (2001) 791–795 H., Tsukamoto, T., Hashimoto, G., Marchesi, E., 1999b. HPLC profiles of floral anthocyanins in the native taxa of Petunia (Solanaceae). Tech. Bull. Fac. Hort. Chiba Univ. 53, 135–144.

Austin, M., Noel, J., 2003. The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Report 20, 79–110.

Bailey JA, Mansfield JW., 1982. Metabolism of phytoalexins. Glasgow, Blackie, p. 334

Balasubramanyam, K., Altaf, M., Varier, RA., Swaminathan, V., Ravindran, A., Sadhale, PP., Kundu, TK., 2004. Polyisoprenylated benzophenone, garcinol, a natural histone acetyltransferase inhibitor, represses chromatin transcription and alters globalgene expression. J Biol Chem, 279(32),33716-26.

Baldwin, I.T., 1989. Mechanism of damage-induced alkaloid production in wild tobacco. J. Chem. Ecol. 15(5), 1661-1680.

Baloglu, E., Kingston, D. G. I., 1999. The structures of over 350 taxane diterpenoids are classified and presented with information on their plant source, yield, melting point, and optical activity. The biotransformations and biosynthesis of the taxoids are also reviewed. J. Nat. Prod. 62, 1448-1472.

Barillas, W., Beerhues, L., 1997. 3-Hydroxybenzoate:coenzyme A ligase and 4-coumarate:coenzyme A ligase from cultured cells of *Centaurium erythraea*. Planta 202, 112 – 116.

Batard, Y., Schalk, M., Pierrel, M., Zimmerlin, A., Durst, F., Werck-Reichhart, D., 1997. Regulation of the cinnamate 4-hydroxylase (CYP73A1) in Jerusalem artichoke tubers in response to wounding and chemical treatments. Plant Physiol. 111**,** 951–959.

Bauer, K., Garbe, D., 1985. Common Fragrance and Flavor Materials: Preparation, Properties and Uses. Deerfield Beach, FL: VCH Publishers.

Bean, W., 1981. Trees and Shrubs Hardy in Great Britain. Vol. 1 - 4 and Supplement. Murray.

Becker-André, M., Schulze-Lefert, P., Hahlbrock, K., 1991. Structural comparison, modes of expression, and putative *cis*-acting elements of the two 4-coumarate:CoA ligase genes in potato. J. Biol. Chem. 266**,** 8551-8559.

Beerhues, L., 1996. Benzophenone synthase from cultured cells of *Centaurium erythraea*. FEBS Lett. 383, 264-266.

Beerhues, L., Berger, U., 1994. Xanthones in cell suspension cultures of two *Centaurium* species. Phytochemistry 35, 1227–1231.

Bell-Lelong, D.A., Cusumano, J.C., Meyer, K., and Chapple, C., 1997. Cinnamate 4-hydroxylase expression in *Arabidopsis.* Plant Physiol. 113**,** 729–738.

Bergstrom, JD., Dufresne, C., Bills, GF., Nallin-Omstead, M., Byrne, K., 1995. Discovery, biosynthesis, and mechanism of action of the zaragozic acids: potent inhibitors of squalene synthase. Annu Rev Microbiol 49**,** 607–639.

Beuerle, T., Pichersky, E., 2002. Purification and characterization of benzoate:Coenzyme A ligase from Clarkia breweri. Arch Biochem Biophys 400:258-264.

Birnboin, H.C., Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nuc. Acid. Res.* 7, 1513−1523.

Bjorklund, J. A., Leete, E., 1992. Biosynthesis of the Benzoyl Moiety of Cocaine from Cinnamic Acid via (R)-(+)-3-Hydroxy-3-phenylpropanoic Acid. Phytochemistry, 31, 3883-3887.

Boari, F., Malone, M., 1993. Wound-Induced hydraulic signals: survey of occurrence in a range of species. J. Exp. Bot. 44, 741–746.

Bolwell, GP., Robbins, MP., Dixon, RA., 1985. Metabolic changes in elicitor-treated bean cells: enzymic responses in relation to rapid changes in cell wall composition. Eur J Biochem.; 148, 571–578.

Bolwell, GP., Cramer, CL., Lamb, CJ., Schuch, W., Dixon, RA., 1986. L-Phenylalanine ammonia-lyase from Phaseolus vulgaris: modulation of the levels of active enzyme by trans-cinnamic acid. Planta.; 169, 97–107.

Booij-James, IS., Dube, SK., Jansen, MA., Edelman M., Mattoo, AK., 2000. Ultraviolet-B radiation impacts light-mediated turnover of the photosystem II reaction center heterodimer in Arabidopsis mutants altered in phenolic metabolism." Plant Physiol 124(3), 1275-84. PMID: 11080303 Acad. Sci. USA, 92, 4076.

Boudet, A. M., 2000. Lignins and lignification: selected issues. Plant Physiol. Biochem.38:8196.

Bradford, J. M., 1976. Partial revision of the Acatia subgenys Acartiura (Copepoda: Calanoida: Acartiidae). N. Z. J. Mar. Freshw. Res., 10, 159–202.

Bradley, D.J., Kjellborn, P., Lamb, C.J., 1992. Elicitor- and wound-induced oxidative cross-linking of a proline-rich plant cell wall protein: a novel, rapid defense response. Cell 70, 21-30.

Bradshaw, H. D. Jr., Stettler, R. F., 1995. Molecular genetics of growth and development in *Populus.* IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics 139**,** 963–973.

Braun, J., Tevini, M., 1993. Regulation of UV-protective pigment synthesis in the epidermal layer of rye seedlings. Photochem Photobiol 57, 318–323.

Brooks, CJW., Watson, DG., 1985. Phytoalexins. Nat Prod Rep 2, 427–459.

Brown, J.E., Rice-Evans, C.A., 1998. Luteolin-rich artichoke extract protects low-density lipoprotein from oxidation in vitro. Free Radical Research 29:247-255.

Brucker, W., 1957. Hemmungsanalytische Studien zur Phenolcarbonsäurebildung durch Phycomyces blakesleeanus. Arch Mikrobiol. 26 (3):302–306.

Buell, C.R., Somerville, S.C., 1995. Expression of defenserelated and putative signaling genes during tolerant and susceptible interactions of *Arabidopsis* with *Xanthomonas campestris* pv. *campestris*. Mol. Plant-Microbe Interact. 8**,** 435–443.

Butland, SL., Chow, ML., Ellis, BE., 1998. A diverse family of phenylalanine ammonia-lyase genes in pine tree and cell cultures. Plant Mol Biol.;37, 15–24

Carnduff, J., Hargreaves, K. R., Nechvatal, A., 1966. Constituents of the *Hypericum androsaemum*. Phytochemistry, 5, 1029.

Celis, H., Romero, I., 1987. The phosphate-pyrophosphate exchange and hydrolytic reactions of the membrane-bound pyrophosphatase of *Rhodospirillum rubrum*: effects of pH and divalent cations. J. Bioenerg. Biomembr. 19, 255-272.

Celis, H., Romero, I., Gómez-Puyou A., 1985. The phosphate-pyrophosphate exchange and hydrolytic reactions of the membrane-bound pyrophosphatase of Rhodospirillum rubrum: effects of Mg2+, phosphate, and pyrophosphate. Arch Biochem Biophys. 1;236(2), 766–774.

Chappell, J., Hahlbrock, K., 1984. Transcription of plant defence genes in response to UV light or fungal elicitor. Nature 311, 76-78.

Chevallier, A., 1996. The Encyclopedia of Medicinal Plants Dorling Kindersley. London ISBN 9-780751-303148.

Chiej, R., 1984. Encyclopaedia of Medicinal Plants. MacDonald ISBN 0-356-10541-5.

Cline, J., Braman, JC., Hogrefe, HH., 1996. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. Nucl. Acids Res. 24, 3546-3551; doi:10.1093/nar/24.18.3546.

Cohen, S.N., Chang, A.C.Y., Hsu, I., 1972. Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. USA. 69, 2110–2114.

Costa, A.F., 1994. Farmacognosia vol. II, Fundação Calouste Gulbenkian, Lisboa. pp. 1022–1023.

Cramer, C.L., Edwards, K., Dron, M., Liang, X., Dildine, S.L., Bolwell, G.P., Dixon, R.A., Lamb, C.J., and Schuch, W., 1989. Phenylalanine ammonia-lyase gene organization and structure. Plant Mol. Biol. 12, 367–383.

Croteau, R., Kutchan, T.M., Lewis, N.G., 2000. Natural products (secondary metabolites). In Biochemistry and Molecular Biology of Plants, B.B. Buchanan, W. Gruissem, and R.L. Jones, eds (Rockville, MD: American Society of Plant Physiologists), pp. 1250–1318.

Cuesta Rubio, O., Cuellar Cuellar, A., Rojas, N., Velez Castro, H., Rastrelli, L., Aquino, R., 1999. A polyisoprenylated benzophenone from Cuban propolis. J. Nat. Prod. 62, 1013-1015.

Cukovic, D., Ehlting, J., VanZiffle, J. and Douglas, C.J., 2001. Structure and evolution of 4-coumarate: coenzyme A ligase (4CL) gene families. Biol. Chem. 382, 645-654.

Darvill, AG, Albersheim, P., 1984. Phytoalexins and their elicitors--a defense against microbial infection in plants. Ann Rev Plant Physiol 35, 243-275.

Davis, P.H. ed., 1967. Flora of Turkey and the East Aegean Islands. Vol. 2:366.

De Luca, V., St Pierre, B., 2000. The cell and developmental biology of alkaloid biosynthesis. Trends Plant Sci. 5, 168-173.

Dewick, PM., 1997. Medicinal natural products. A biosynthetic approach. Wiley, Chichester.

Dias, A. C. P., Seabra, R. M., Andrade, P. B., Fernandes-Ferreira, M., 1999. The development and evaluation of an HPLC-DAD method for the analysis of the phenolic fractions from in vivo and in vitro biomass of *Hypericum* species. J. Liq. Chromatogr. Relat. Technol., 22 (2), 215-227.

Dias, A. C. P., Seabra, R. M., Andrade, P. B., Fernandes-Ferreira, M., 2000. Xanthone biosynthesis and accumulation in *calli* and suspended cells of *Hypericum androsaemum*. Plant Sci., 150, 93-101.

Dieffenbach, CW., Lowe, TMJ., Dveksler, GS., 1995. General Concepts for PCR Primer Design, Cold Spring Harbor Laboratory Press, New York, 133-155.

Dignum, M.J.W., Kerler J., Verpoorte R., 2001. Vanilla production: technological, chemical, and biosynthetic aspects. Food Res. Int. 17, 199–219.

Diógenes Aparício Garcia Cortez; Benício Alves Abreu Filho; Celso Vataru Nakamura; Benedito Prado Dias Filho; Andrew Marston; Kurt Hostettmann., 2002. Antibacterial Activity of a Biphenyl and Xanthones from *Kielmeyera coriacea* Pharmaceutical Biology, Vol. 40, No. 7, pp. 485–489.

Dirr, M. A., Heuser, M. W., 1987. The Reference Manual of Woody Plant Propagation. Athens Ga. Varsity Press ISBN 0942375009.

Dixon, R.A., Paiva, N.L., 1995. Stress-induced phenylpropanoid metabolism. Plant Cell 7, 1085–1097.

Dixon RA., Achnine L., Kota P., Liu C-J., Reddy MSS., Wang L., 2002a. The phenylpropanoid pathway and plant defence- a genomics perspective. Molecular Plant Pathology, 3(5), 371-390.

Douglas, C. J., 1996. Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. Trends Plant Sci. 1**,** 171-178.

Douglas, C., Hoffmann, H., Schulz, W., Hahlbrock, K., 1987. Structure and elicitor- or UV-light-stimulated expression of two 4-coumarate coenzyme A ligase genes in parsley. EMBO J 6, 1189-1196.

Draper, J., 1997. Salicylate, superoxide synthesis and cell suicide in plant defense. Trends Plant Sci. 2, 162 - 165.

Dudareva, N., D'Auria, JC., Nam, KH., Raguso, RA., Pichersky, E., 1998. Acetyl-CoA:benzyl alcohol acetyltransferase: an enzyme involved in floral scent production in *Clarkia breweri*. Plant J 14, 297–304.

Dudareva, N., Pichersky, E., 2000. Biochemical and molecular genetic aspects of floral scents. Plant Physiol 122(3); 627-33. PMID: 10712525.

Dudareva, N., Pichersky, E., Gershenzon, J., 2004. Biochemistry of plant volatiles. Plant Physiol 135(4), 1893-902. PMID: 15326281.

Ecker, JR., Davis, RW., 1987. Plant defence genes are regulated by ethylene. Proc Natl Acad Sci USA 84, 5202–5206.

Egland, PG., Gibson, J., Harwood, CS., 1995. Benzoate-coenzyme A ligase, encoded by badA, is one of three ligases able to catalyze benzoyl-coenzyme A formation during anaerobic growth of Rhodopseudomonas palustris on benzoate. J Bacteriol. Nov; 177(22), 6545–6551.

Ehlting, J., BuÈ ttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., Kombrink, E., 1999. Three 4-coumarate: coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. Plant J. 19, 9-20.

Engler, M.J., Richardson, C.C., 1982. DNA ligases. Enzymes, 15, 3–29.

Fahrendorf, T., Dixon, R.A., 1993. Stress responses in alfalfa (*Medicago sativa* L.). XVIII. Molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450. Arch. Biochem. Biophys. 305**,** 509–515.

Fett-Neto, AG., Zhang, WY., DiCosmo, F., 1994. Kinetics of taxol production, growth and nutrient uptake in cell suspensions of *Taxus cuspidata*. Biotechnol Bioeng 44, 205–210.

Fleming, PE., Mocek, U., Floss, HG., 1993. Biosynthesis of taxoids and mode of formation of the taxol side chain. J Am Chem Soc 115, 805–807.

Fliegmann, J., Schroder, G., Schanz, S., Britsch, L., Schroder, J., 1992. Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (Pinus sylvestris), and differential regulation of these and related enzyme activities in stressed plants." Plant Mol Biol 18(3), 489-503. PMID: 1536925.

Flora Europea, 1968. vol II. Cambridge University Press, London, p. 263.

Florian, G., Paulo, F., Kenjiro, T., Fumi, T., Norio, S., Raymond B., 2001. Influence of trans-cis isomerisation of coumaric acid substituents on colour variance and stabilisation in anthocyanins. Phytochemistry 57 791–795.

French, CJ., Vance, CP., Towers, GHN., 1976. Conversion of p-coumaric acid to p-hydroxybenzoic acid by cell free extracts of potato tubers and Polyporus hispidus. Phytochemistry 15, 564-566.

Fukasawa-Akada, T., Kung, S., and Watson, J.C., 1996. Phenylalanine ammonia-lyase gene structure, expression, and evolution in *Nicotiana*. Plant Mol. Biol. 30**,** 711–722.

Fulda, M., Heinz, E., Wolter, F. P., 1994. The *fadD* gene of *Escherichia coli* K-12 is located close to *rnd* at 39.6 min of the chromosomal map and is a new member of the AMP-binding protein family. Mol. Gen. Genet. 242**,** 241-249.

Fuller, R.W., Blunt, J.W., Boswell, J.L., Cardellina II, J.H., Boyd, M.R., 1999. Guttiferone F, the first prenylated benzophenone from Allanblackia stuhlmannii. J. Nat. Prod.,62, 130-2.

Funk, C., Brodelius, E., 1990a. Phenylpropanoid Metabolism in Suspension Cultures of Vanilla planifolia Andr. : II. Effects of Precursor Feeding and Metabolic Inhibitors. Plant Physiol. Sep;94(1), 95–101.

Funk, C., Brodelius, E., 1990b. Phenylpropanoid Metabolism in Suspension Cultures of Vanilla planifolia Andr. : III. Conversion of 4-Methoxycinnamic Acids into 4-Hydroxybenzoic Acids. Plant Physiol. Sep;94(1), 102–108.

Gasson, MJ., Kitamura, Y., McLauchlan, WR., Narbad, A., Parr, AJ., Parsons, ELH., Payne, J., Rhodes, MJC., Walton, NJ., 1998. Metabolism of ferulic acid to vanillin. A bacterial gene of the enoyl-SCoA hydratase/isomerase superfamily encodes an enzyme for the hydration and cleavage of a hydroxycinnamic acid SCoA thioester. J Biol Chem 273, 4163–4170.

Goldspiel, B. R., 1997. Clinical overview of the taxanes. Pharmacotherapy 17, 1105–1255.

Grand, C., Boudet, A., Boudet, AM., 1983. Isoenzymes of hydroxycinnamate:CoA ligase from poplar stems, properties, and tissue distribution. Planta 159, 225–229.

Grieve, 1984. A Modern Herbal. Penguin ISBN 0-14-046-440-9.

Gundlach, H., Muller, M. J., Kutchan, M. J., Zenk, M. H., 1992. JA is a signal transducer elicitor-induced plant cell cultures. Proceedings of the National Academy of Sciences USA 89, 2389-2393.

Habereder, H., Schröder, G., Ebel, J., 1989. Rapid induction of phenylalanine ammonia-lyase and chalcone synthase mRNAs during fungus infection of soybean (*Glycine max* L.) roots or elicitor treatment of soybean cell cultures at the onset of phytoalexin synthesis. Planta 177, 58–65.

Hahlbrock, K., Knobloch, KH., Kreuzaler, F., Potts, JRM., Wellmann, E., 1976. Coordinated induction and subsequent activity changes of two groups of metabolically interrelated enzymes. Eur J Biochem.; 61, 199–206.

Hahlbrock, K., Grisebach, H., 1979. Enzymatic controls in the biosynthesis of lignin and flavonoids. Annu Rev Plant Phys 30, 105-136.

Hahlbrock, K., Scheel, D., 1989. Physiology and molecular biology of phenylpropanoid metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol., 40, 347-369.

Hamberger, B., Hahlbrock, K. 2004. The 4-coumarate : CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. Proceedings of the National Academy of Sciences of the United States of America 101, 7, 2209-2214.

Hammerschmidt, R., 1999. Induced disease resistance: How do induced plants stop pathogens? Physiol. Mol. Plant Pathol. 55, 77–84.

Harborne, J. B.,Williams, C. A., 2000. Advances in flavonoid research since 1992. Phytochemistry 55, 481.

Harding, SA., Leshkeich, J., Chiang, VL., Tsai, CJ., 2002. Differential substrate inhibition couples kinetically distinct 4-coumarate: conezyme A ligases with spatially distinct metabolic roles in quaking aspen. Plant Physiol 128, 428–438.

Haskins, FA., Williams, LG., Gorz, HJ., 1964. Light-induced *trans* to- *cis* conversion of β-*D*-glucosyl *o*-hydroxycinnamic acid in *Melilotus alba* leaves. Plant Physiol 39, 777–781.

Havkin-Frenkel, D., Podstolski, A., Knorr, D., 1996. Effect of light on vanillin precursors formation. b in vitro cultures of *Vanilla planifolia*. Plant Cell Tissue Organ. Cult. 45, 133– 136.

Havkin-Frenkel, D., Podstolski, A., Witkowska, E., Molecki, P., Mikolajczyk, M., 1999. Vanillin biosynthetic pathways, an overview. In: Fu, T.J., Singh, G., Curtis, W.R., eds. Plant cell and tissue culture for the production of food ingredients. Kluwer Academic Press/Plenum Publishers, New York, pp. 35–43.

Heide, L., Tabata, M., 1987. Geranylpyrophosphate: *p*-Hydroxybenzoate geranyltransferase activity in extracts of *Lithospermum erythrorhizon* cell cultures. Phytochemistry 26, 1651–1655.

Hibi, N., Higashiguchi, S., Hashimoto, T., Yamada, Y., 1994. Gene expression in tobacco low-nicotine mutants. Plant Cell 6, 723–735.

Holub, J., Hanus, J., Hanke, D.E., Strnad, M., 1998. Biological activity of cytokinins derived from ortho- and meta-hydroxybenzyladenine. Plant Growth Regul. 26, 109–115.

Hopper, D. J., 1995. Tropine dehydrogenase: purification, some properties and an evaluation of its role in the bacterial metabolism of tropine. Biochem. J. 307**,** 603–608.

Hosokawa, K., Fukunaga, Y., Fukushi, E., Kawabata, J., 1995a. Seven acylated anthocyanins in the blue flowers of Hyacinthus orientalis. Phytochemistry 38, 1293–1298.

Hosokawa, K., Fukunaga, Y., Fukushi, E., Kawabata, J., 1995b. Acylated anthocyanins from red Hyacinthus orientalis. Phytochemistry 39, 1437–1441.

Hrazdina, G., 2003. Response of scab-susceptible (McIntosh) and scab-resistant (Liberty) apple tissues to treatment with yeast extract and Venturia inaequalis. Phytochemistry. 64(2), 485-492.

Hu, L.H., Sim, K.Y., 2000. Sampsoniones A–M, a unique family of caged polyprenylated benzoylphloroglucinol derivatives, from Hypericum sampsonii. Tetrahedron 56, 1379–1386.

Hu, L.H., Sim, K.Y., 1999. Cytotoxic polyprenylated benzoylphloroglucinol derivatives with an unusual adamantyl skeleton from Hypericum sampsonii (Guttiferae). Org. Lett., 23, 879.

Hu, W.J., Kawaoka, A., Tsai, C.J., Lung, J.H., Osakabe, K., Ebinuma, H., Chiang, V.L., 1998. Compartmentalized expression of two structurally and functionally distinct 4-coumarate-CoA ligase genes in aspen (Populus tremuloides). Proc. Nat. Acad. Sci. USA, 95, 5407-5412.

Hu, W.J., Lung, J., Harding, S.A., Popko, J.L., Ralph, J., Stokke, D.D., Tsai, C.-J., Chiang, V.L., 1999. Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat. Biotechnol. 17, 808–812.

Humphreys, JM., Hemm, MR., Chapple, C., 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci USA 96, 10045–10050.

Huxley, A., 1992. The New RHS Dictionary of Gardening. MacMillan Press ISBN 0-333-474945.

Jameson, PE.,1994. Cytokinin metabolism and compartmentation. In: Mok DWS, Mok MC (eds) Cytokinins: Chemistry, Activity and Function. CRC Press, Boca Raton, London, Tokyo, pp 113–128.

Junghanns, KT., Kneusel, RE., Baumert, A., Maier, W., Gröger, D., Matern, U., 1995. Molecular cloning and heterologous expression of acridone synthase from elicited *Ruta graveolens* L. cell suspension cultures. Plant Mol Biol 27, 681–692.

Kahnt, G., 1967. Trans-cis-equilibrium of hydroxycinnamic acids during irradiation of aqueous solutions at different pH. Phytochemistry 6, 755–758.

Kato, T., Ito, S., Fujita, K., 1986. Tyrosinase-catalysed binding of 3,4 dihydroxyphenylalanine with proteins through the sulfydryl group. Biochim Biophys Acta 881, 415–421.

Katsuyoshi, H., Tomoaki, N., Kazuchika, Y., Kazuhiko, F., Ryuichiro, K., Yuji, T., 2003. 4-Coumarate: coenzyme A ligase in black locust (Robinia pseudoacacia)catalyses the conversion of sinapate to sinapoyl-CoA J Plant Res (2004) 117, 303–310.

Kitanov, G. M., Nedialkov, P. T., 1998. Mangiferin and isomangiferin in some *Hypericum* species. Biochem. Syst. Ecol., 26, 647-653.

Koch, MA., Weisshaar, B., Kroymann, J., Haubold, B., Mitchell-Olds, T., 2001. Comparative genomics and regulatory evolution: conservation and function of the Chs and Apetala3 promoters. Mol Biol Evol 18(10); 1882-91. PMID, 11557794.

Kodan, A., Kuroda, H., Sakai, F., 2002. A stilbene synthase from Japanese red pine (*Pinus densiflora*): implications for phytoalexin accumulation and down-regulation of flavonoid biosynthesis. Proc Natl Acad Sci U S A 99(5); 3335-9. PMID, 11880657.

Koes, RE., Spelt, CE., van den Elzen, PJ., Mol, JN., 1989. Cloning and molecular characterization of the chalcone synthase multigene family of Petunia hybrida. Gene 81(2), 245-57. PMID: 2806915.

Komarov, V. L., 1968. Flora of the USSR. Israel Program for Scientific Translation.

Kondo, T., Yoshida, K., Yoshikane, M., Goto, T., 1991. Mechanism for color development in purple flower of Commelina communis. Agric. Biol. Chem. 55, 2919–2921.

Kokubun, T., Harborne, JB., 1995a. Antifungal biphenyl compounds are the phytoalexins of the sapwood of *Sorbus aucuparia*. Phytochemistry 40, 57–59.

Kokubun, T., Harborne, JB., 1995b. Phytoalexin induction in the sapwood of plants of the Maloideae (Rosaceae): Biphenyls and dibenzofurans. Phytochemistry 40, 1649–1654.

Knobloch, K. H., Hahlbrock, K., 1975. Isoenzymes of *p*-coumarate:CoA ligase from cell suspension cultures of *Glycine max*. Eur. J. Biochem. 52, 311-320.

Knobloch, K-H., Hahlbrock, K., 1975. Isoenzymes of *p*-coumarate:CoA ligase from cell suspension cultures of *Glycine max* (soybeans). Eur J Biochem 52, 311-320.

Knobloch, K-H., Hahlbrock, K., 1977. 4-Coumarate:CoA ligase from cell suspension cultures of *Petroselinum hortense* Hoffm. Arch Biochem Biophys 184, 237-248.

Kumar, A., Ellis, B., 2003. 4-Coumarate:CoA ligase gene family in *Rubus idaeus*: cDNA structures,evolution, and expression. Plant Molecular Biology 31, 327–340.

Kutsuki, H., Higuchi, T.,1981. Activities of some enzymes of lignin formation in reaction wood of *Thuja orientalis, Metasequoia glyptostroboides* and *Robinia pseudoacacia*. Planta 152, 365–368.

Kutsuki, H., Shimada, M., Higuchi, T., 1982. Regulatory role of cinnamyl alcohol dehydrogenase in the formation of guaiacyl and syringyl lignins. Phytochemistry 21, 19–23.

Kutsuki, H., Shimada, M., Higuchi, T., 1982. Distribution and roles of *p*-hydroxycinnamate: CoA ligase in lignin biosynthesis. Phytochem 21, 267–271.

Laempe, D., Jahn, M., Breese, K., Fuchs, G., 2001. Anaerobic Metabolism of 3-Hydroxybenzoate by the Denitrifying Bacterium *Thauera aromatica.* Journal of Bacteriology, February 2001, p. 968-979, Vol. 183, No. 3.

Laemmli, UK., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lamb, CJ., 1979. Regulation of enzyme levels in phenylpropanoid biosynthesis: characterization of the modulation by light and pathway intermediates. Arch Biochem Biophys.; 192, 311–317.

Lange, BM., Wildung, MR., McCaskill, D., Croteau, R., 1998. A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway. Proc Natl Acad Sci USA.; 95, 2100–2104.

Launert, **E., 1981.** The Hamlyn Guide to Edible and Medicinal Plants of Europe and. Northern Europe., 4th. impression 1989. Hamlyn. ISBN No.0-600-56395-2.

Lee, D., Douglas, CJ., 1996. Two divergent members of a tobacco (*Nicotiana tabacum*) 4-coumarate:CoA ligase (*4CL*) gene family: cDNA structure, gene inheritance and expression, and properties of recombinant proteins. Plant Physiol 112: 193-205.

Lee, H., J. León, Raskin, I., 1995. Biosynthesis and metabolism of salicylic acid. Proc. Natl. Acad. Sci. USA 92, 4076-4079.

Leete, E., Bodem, G. B., 1966. The Biosynthesis of 3-(dimethylamino)- 3-phenylpropanoic acid in yew. Tetrahedron Lett.**,** 33, 3925-3927.

Lehman, I.R., 1974. DNA ligase: structure, mechanism, and function**.**Science, 186, 790–797.

León, J., Yalpani, N., Raskin, I., Lawton, M. A., 1993. Induction of benzoic acid 2 hydroxylase in virus-inoculated tobacco. Plant Physiol. 103, 323-328.

Li, L., Popko, JL., Umezawa, T., Chiang, VL., 2000. 5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. J Biol Chem 275, 6537–6545.

Lindermayr, C., Mollers, B., Fliegmann, J., Uhlmann, A., Lottspeich, F., Meimberg, H., Ebel, J., 2002. Divergent members of a soybean (*Glycine max* L.) 4-coumarate: coenzyme A ligase gene family. Primary structures, catalytic properties, and differential expression. Eur J Biochem 269, 1304–1315.

LINSMAIER, E. M., SKOOG, F., 1965. Organic growth factor requirements of tobacco tissue cultures. Physiol. Plantarum 18, 100-27.

Liu, B., Falkenstein-Paul, H., Schmidt, W., Beerhues, L., 2003. Benzophenone synthase and chalcone synthase from *Hypericum androsaemum* cell cultures: cDNA cloning, functional expression, and site-directed mutagenesis of two polyketide synthases. The Plant Journal 34, 847-855.

Liu, B., Beuerle, T., Klundt, T., Beerhues, L., 2004. Biphenyl synthase from yeast-extract-treated cell cultures of *Sorbus aucuparia*. Planta 218, 492-496.

Liu, B.; Raeth, T.; Beuerle, T.; Beerhues, L. 2006. planta submitted.

Loake, G., Choudhary, AD., Harrison, MJ., Mavandad, M., Lamb, CJ., Dixon, RA., 1991. Phenylpropanoid pathway intermediates regulate transient expression of a chalcone synthase gene promoter in electroporated protoplasts. Plant Cell; 3, 829–840.

Locher, R., Martin, V., Grison, R., Pilet, PE., 1994. Cell wallbound *trans*- and *cis*-ferulic acids in growing maize roots. Physiol Plant 90, 734–738.

Logemann, E., Parniske, M., Hahlbrock, K., 1995. Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proc Natl Acad Sci USA 92, 5905–5909.

Lorenzen, M., Racicot, V., Strack, D., Chapple, C., 1996. Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of arabidopsis. Plant Physiol 112(4), 1625-30. PMID: 8972602.

Loscher, R., Heide, L., 1994. Biosynthesis of p-hydroxybenzoate from p-coumarate and p-coumaroyl-coenzyme A in cell-free extracts of *Lithospermum erythrorhizon* cell cultures. Plant Physiol, 106, 271-279.

Lust, J., 1983. The Herb Book. Bantam books ISBN 0-553-23827-2 Lots of information tightly crammed into a fairly small book.

Lozoya, E., Hoffmann, H., Douglas, C.J., Schulz, W., Scheel, D., Hahlbrock, K., 1988. Primary structures and catalytic properties of isoenzymes encoded by the two 4-coumarate: CoA ligase genes in parsley. Eur. J. Biochem. 176, 661- 667.

Lüderitz, T., Schatz, G., Grisebach, H., 1982. Enzymic synthesis of lignin precursors. Purification and properties of 4-coumarate:CoA ligase from cambial sap of spruce (*Picea abies L.*) European Journal Of Biochemistry / FEBS Volume 123, Issue 3 , April, Pages 583-586 ISSN: 0014-2956.

Madigan, M. T., Martinko J. M., Parker, J., 1997. Brock Biology of Microorganisms. Prentice Hall, Upper Saddle River, N.J.

Malamy, J., Carr, JP., Klessig, DF., Raskin, I., 1990. Salicylic acid: a likely endogenous signal in the resistance response of tobacco to tobacco mosaic virus. Science 250, 1002-1004.

Maruyama, K., Sugano, S., 1994. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. Gene 138, 171–174.

Mathis, C., Ourisson, G., 1964. Étude chimio-taxonomique du genre *Hypericum* III. Répartition des carbures saturés et des monoterpènes dans les huiles essentielles d'*Hypericum*. *Phytochemistry*, 3, 133-141.

Mathis, C., Ourisson, G., 1964. Etude chimio-taxonomique du genre *Hypericum*-IV. Repartition des sesquiterpenes, des alcools monoterpeniques et des aldehydes satures dans les huilles essentielles d'*Hypericum*. Phytochemistry, 3, 377-378.

Matsui, N., Fukushima, K., Kamada, K., Nishikawa, Y., Yasuda, S., 1994. On the behavior of monolignol glucosides. I. Synthesis of monolignol glucosides labeled with ^{2}H at the hydroxymethyl group of side chain, and polymerization of the labeled monolignol in vitro. Holzforschung 48, 215–221.

Matsumoto, K., Akao, Y., Kobayashi, E., Ito, T., Ohguchi, K., Tanaka, T., Iinuma, M., Nozawa,Y., 2003. Cytotoxic benzophenone derivatives from Garcinia species display a strong apoptosis-inducing effect against human leukemia cell lines. Biol. Pharm. Bull., 26, 569.

Mavandad, M., Edwards, R., Liang, X., Lamb, CJ., Dixon, RA., 1990. Effects of trans-cinnamic acid on expression of the bean phenylalanine ammonia-lyase gene family. Plant Physiol.; 94, 671–680.

Meng, H., Campbell, WH., 1997. Facile enzymatic synthesis of caffeoyl CoA. Phytochemistry.;44, 605–608.

Metraux, JP., Signer, H., Ryals, J., Ward, E., Wyss-Benz, M., Gaudin, J., Raschdorf, K., Blum, W., Inverardi, B., 1990. Increase in salicylic acid at the onset of systemic acquired resistance in cucumber. Science 250, 1004-1006.

Meuwly, P., Molders, W., Buchala, A., Metraux, JP., 1995. Local and systemic biosynthesis of salicylic-acid in infected cucumber plants. Plant Physiol. 109:1107–14 Annu. Rev. Phytopathol. 2005.43, 545-580.

Moniz, de Sa M., Subramaniam, R., Williams, FE., Douglas, CJ., 1992. Rapid activation of phenylpropanoid metabolism in elicitor-treated hybrid poplar (*Populus trichocarpa* Torr and Gray _ *Populus deltoides* Marsh) suspension-cultured cells. Plant Physiol 98, 728–737.

Moore, B. S., Hertweck, C., Hopke, J. N., and other authors 2002. Plant-like biosynthetic pathways in bacteria: from benzoic acid to chalcone. *J Nat Prod* 65, 1956–1962.

Nagai, N., Kitauchi, F., Shimosaka, M., Okazaki, M., 1994. Cloning and sequencing of a full length cDNA coding for phenylalanine ammonia-lyase from tobacco cell culture. Plant Physiol. 104, 1091–1092.

Ni, W., Fahrendorf, T., Ballance, GM., Lamb, CJ., Dixon, RA., 1996. Stress responses in alfalfa (*Medicago sativa L.*): XX. Transcriptional activation of phenylpropanoid pathway genes in elicitor-treated cell suspension cultures. Plant Mol Biol.; 30, 427–438.

Nielsen, H., Arends, P., 1979. Xanthone constituents of *Hypericum androsaemum*. J. Nat. Prod., 42, 301-304.

Nogueira, T., Duarte, F., Venâncio, F., Tavares, R., Lousã, M., Bicchi, C., Rubiolo, P.,1998. Aspectos quimiotaxonomicos do género *Hypericum* L. em Portugal. Silva Lusitanica, 6 (1), 55-61.

Ohashi, H., Yamamoto, E., Lewis, NG., Towers, GHN., 1987. 5- Hydroxyferulic acid in *Zea mays* and *Hordeum vulgare* cell walls. Phytochemistry 26, 1915–1916.

Osakabe, K., Tsao, CC., Li, L., Popko, JL., Umezawa, T., Carraway, DT., Smeltzer, RH., Joshi, CP., Chiang, VL., 1999. Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proc Natl Acad Sci USA 96, 8955–8960.

Parker, JE., Schulte, W., Hahlbrock, K., Scheel, D.,1991. An extracellular glycoprotein from *Phytophthora megasperma* f. sp. *glycinea* elicits phytoalexin synthesis in cultured parsley cells and protoplasts. Mol Plant-Microbe Interact, 4, 19–27.

Pedras, MSC., Khan, AQ., 2000. Biotransformation of the phytoalexin camalexin by the phytopathogen *Rhizoctonia solani.* Phytochemistry 53, 59–69.

Pedras, MSC., Okanga, FI., Zaharia, IL., Khan, AQ., 2000. Phytoalexins from crucifers synthesis, biosynthesis, and biotransformation. Phytochemistry 53, 161–176.

Peters, S., Schmidt, W., Beerhues, L., 1998. Regioselective oxidative phenol couplings of 2,3',4,6-tetrahydroxybenzophenone in cell cultures of *Centaurium erythraea* RAFN and *Hypericum androsaemum* L. Planta, 204, 64-69.

Peterson, M., Strack, D., Matern, U., 1999. Biosynthesis of phenylpropanoids and related compounds. In Biochemistry of Plant Secondary Metabolism. Edited by Wink, M. pp. 151–221. Sheffield Academic Press, Sheffield.

Pieterse, CMJ., Van Wees, SCM., Ton, J., Van, Pelt, JA., van Loon, LC., 2002. Signalling in rhizobacteria-induced systemic resistance in *Arabidopsis thaliana*. Plant Biol. 4, 535–44.

Podstolski, A., Havkin-Frenkel, D., Malinowski, J., Blount, J., Kourteva, G., Dixon, R.A., 2002. Unusual 4-hydroxybenzaldehyde synthase activity from tissue cultures of the vanilla orchid *Vanilla planifolia*. Phytochemistry 61, 611– 620.

Preisig-Mueller, R., Gnau, P., Kindl, H., 1995. The inducible 9,10 dihydrophenanthrene pathway: characterization and expression of bibenzyl synthase and *S*adenosylhomocysteine hydrolase. Arch. Biochem. Biophys. 317, 201–207.

Randahl, H., 1979. Characterization of the membrane-bound inorganic pyrophosphatase in Rhodospirillum rubrum. Eur J Biochem. Dec;102(1), 251–256.

Rasmussen, S., Rudolph, H. 1997. Isolation, purification and characterization of UDP-glucose: *cis-p*-coumaric acid β-*D*glucosyltransferase from *Sphagnum fallax*. Phytochemistry 46, 449–453.

Ribnicky, DM., Shulaev, V., Raskin, I., 1998. Intermediates of salicylic acid biosynthesis in tobacco. Plant Physiol. 118, 565-572.

Russell, D.W., Conn, E.E., 1967. The cinnamic acid 4-hydroxylase of pea seedlings. Arch. Biochem. Biophys. 122, 256–258.

Sarma, AD., Sreelakshmi, Y., Sharma, R., 1998. Differential expression and properties of phenylalanine ammonia-lyase isoforms in tomato leaves. Phytochemistry 49, 2233–2243.

Schaefer, B.C., 1995. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. Anal. Biochem., 227, 255–273.

Schmidt, W., Beerhues, L., 1997. Alternative pathways of xanthone biosynthesis in cell cultures of *Hypericum androsaemum* L. FEBS Letters 420, 143 –146.

Schmidt, W., Peters, S., Beerhues, L., 2000. Xanthone 6-hydroxylase from cell cultures of *Centaurium erythraea* RAFN and *Hypericum androsaemum L.* Phytochemistry 53, 427-431.

Schneider, K., Hovel, K., Witzel, K., Hamberger, B., Schomburg, D., Kombrink, E., Stuible, HP., 2003. The substrate specificity-determining amino acid code of 4-coumarate:CoA ligase. Proc Natl Acad Sci USA 100, 8601–8606.

Schnitzler, J-P., Madlung, J., Rose, A., Seitz, HU., 1992. Biosynthesis of p-hydroxybenzoic acid in elicitor-treated carrot cell cultures. Planta, 188, 594-600.

Schroeder, G., Brown, J.W.S., Schroeder, J., 1988. Molecular analysis of resveratrol synthase: cDNA, genomic clones and relationship with chalcone synthase. Eur. J. Biochem. 172, 161–169.

Schröder, J., Raiber, S., Berger, T., Schmidt, A., Schmidt, J., Soares- Sello, A.M., Bardshiri, E., Strack, D., Simpson, T.J., Veit, M., Schro¨ der, G., 1998. Plant polyketide synthases: a chalcone synthasetype enzyme which performs a condensation reaction with methylmalonyl- CoA in the biosynthesis of C-methylated chalcones. Biochemistry 37, 8417–8425.

Seabra, R. M., Vasconcelos, M. H., 1992. Análise de amostras de hipericão existentes no mercado. Rev. Portuguesa Farmacognosia, 17 (3).

Seskar, M., Shulaev, V., Raskin, I., 1998. Endogenous methyl salicylate in pathogen-inoculated tobacco plants. Plant Physiol. 116, 387-392.

Shah, J., Klessig, D.F., 1999. Salicylic acid Signal perception and transduction. In "Biochemistry and Molecular Biology of Plant Hormones", Vol 33, pp 513-541, ed. Hooykaas, P.J.J., M. Hall, and K. Libbenga. Elseiver, UK.

Shi, H., Noguchi, N., Niki, E., 2001. Introducing natural antioxidants In: J. Pokorny, N. Yanishlieva and M. Gordon, Editors, Antioxidants and Food Practical Applications, Woodhead Publishing Limited, Cambridge.

Shi, H., Noguchi, N., Niki, E., 2001. Flavonoids and other polyphenols. Methods Enzymol. 335, 157-166.

Shinshi, H., Mohnen, D., Meins, Jr., 1987. Regulation of a plant pathogenesis-related enzyme inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin. Proc Natl Acad Sci USA 84, 89–93.

Shulaev, V., Silverman, P., Raskin, I., 1997. Airborne signaling by methyl salicylate in plant pathogen resistance.Nature (1997), 385, 718-721.

Shulaev, V., Silverman, P., Raskin, I., 1997. Airborne signaling by methyl salicylate in plant pathogen resistance. Nature 385, 718-721.

Slimestad, R., Aaberg, A., Andersen, M., 1999. Acylated anthocyanins from petunia flowers. Phytochemistry 50, 1081–1086.

Srinivasan, V., Veeresham, C., Bringi, V., Shuler, ML., 1996. Metabolic inhibitors, elicitors and precursors and tool for the probing pathway compartmentation and rate limitation in taxane bio-synthesis. Biotechnol Prog 12, 457–465.

Stuible, H., Buttner, D., Ehlting, J., Hahlbrock, K., Kombrink, E., 2000. Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes. FEBS Lett. 467, 117–122.

Teutsch, H.G., Hasenfratz, M.P., Lesot, A., Stoltz, C., Garnier, J.M., Jeltsch, J.M., Durst, F., Werck-Reichhart, D., 1993. Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. Proc. Natl. Acad. Sci. USA 90, 4102–4106.

Toki, K., Saito, N., Honda, T., 1998. Acylated anthocyanins from the blue-purple flowers of Triteleia bridgesii. Phytochemistry 48, 729– 732.

Towers, G.H.N., Abeysekera, B., 1984. Cell-wall hydroxycinnamate esters as UV-A receptores in phototropic responsis of higher plants- a new hypothesis. Phytochemistry 23 951 – 952.

Triska, Dr., 1975. Hamlyn Encyclopaedia of Plants. Hamlyn ISBN 0-600-33545-3.

Uhlmann, A., Ebel, J., 1993. Molecular cloning and expression of 4-coumarate: coenzyme A ligase, an enzyme involved in the resistance of soybean (Glycine max) against pathoghen infection. Plant Physiol. 102, 1147-1156.

Usher, G., 1974. A dictionary of plants used by man. Constable ISBN 0094579202.

Valentão, P., Andrade, PB., Areias, F., Ferreres, F., Seabra, RM., 1999. Analysis of vervain flavonoids by HPLC/diode array detector method. Its application to quality control. J Agric Food Chem 47, 4579–4582.

Valentão, P., Andrade, PB., Silva, E., Vicente, A., Santos, H., Bastos, ML., Seabra, RM., 2002. Methoxylated xanthones in the quality control of small centaury (*Centaurium erythraea*) flowering tops. J Agric Food Chem 50, 460– 463.

Valentão, P., Dias, A., Ferreira, M., Silva, B., Andrade, P.B., Bastos, M.L., Seabra, R.M., 2003. Variability in phenolic composition of *Hypericum androsaemum*, Natural Product Research 17, pp. 135–140.

Valentão, P., Fernandes, E., Carvalho, F., Andrade, P.B., Seabra, R.M., Bastos, M.L., 2002. Antioxidant activity of *Hypericum androsaemum* infusion: scavenging activity against superoxide radical, hydroxyl radical and hypochlorous acid, Biological and Pharmaceutical Bulletin 25, pp. 1320–1323.

Voo, KS., Whetten, RW., O'Malley, DM., Sederoff, RR., 1995. 4-Coumarate:coenzyme A ligase from loblolly pine xylem. Isolation, characterization, and complementary DNA cloning. Plant Physiol.; 108, 85–97.

Walker, K., Croteau, R., 2001. Taxol biosynthetic genes. Phytochemistry, 58, 1-7.

Wallis, PJ., Rhodes, MJC., 1977. Multiple forms of hydroxycinnamate:CoA ligase in etiolated pea seedlings. Phytochemistry.; 16, 1891–1894.

Wani, M. C., Taylor, H. L., Wall, M. E., Coggen, P., McPhail, A. T., 1971. Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. Journal of the American Chemical Society 93, 2325/2327.

Wanner, L.A., Li, G., Ware, D., Somssich, I.E., Davis, K.R., 1995. The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. Plant Mol. Biol. 27, 327–338.

Ward, E.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahl-Goy, P., Metroux, J.-P., Ryals, J.A., 1991. Coordinate gene activity in response to agents that induce systemic acquired resistance. Plant Cell 3, 1085–1094.

Weihs U., 1995.. Die Eberesche – eine ökologisch wertvolle Baumart bei der Walderneuerung auf Problemstandorten. In : LÖBF (Hrsg.): Weichhölzer und Sukzessionsdynamik in der

naturnahen Waldwirtschaft. Schriftenreihe der Landesanstalt für Ökologie, Bodenordnung und Forsten/Landesamt für Agrarordnung.

Weisshaar, B., G. I. Jenkins., 1998. Phenylpropanoid biosynthesis and its regulation. Curr. Opin. Plant Biol. 1, 251-257

Werner, I., Bacher, A., Eisenreich, W., 1997. Retrobiosynthetic NMR studies with ^{13}C-labeled glucose. Formation of gallic acid in plants and fungi. J Biol Chem 272, 25474 – 25482.

White, R F., 1979. Acetysalicylic acid induces resistance to TMV in tobacco. Virology, 99, 410-412.

Whetten, RW., Sederoff, RR., 1992. Phenylalanine ammonia-lyase from loblolly pine - purification of the enzyme and isolation of complementary DNA clones. Plant Physiol 98, 380-386

Wildon, DC., Thain, JF., Minchin, PEH., Gubb, IR., Reilly, AJ., Skipper, YD., Doherty, HM., O'Donnell, PJ., Bowles, DJ., 1992. Electrical signalling and systemic proteinase inhibitor induction in the wounded plant. Nature 369, 62–65.

Wilks, A. F., 1989. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. *Proc.* Natl. Acad. Sci. USA 86, 1603-1607.

Winkel, BS., 2004. Metabolic channeling in plants. Annu Rev Plant Biol 55, 85-107. PMID: 15725058.

Wu, H., Haig, T., Pratley, J., Lemerle, D., An, M., 2001. Allelochemicals in wheat (*Triticum aestivum* L.): Variation of phenolic acids in shoot tissues. J Chem Ecol 27, 125–135.

Wu, SC., Hahlbrock, K., 1992. In situ localization of phenylpropanoid-related gene expression in different tissues of light-grown and dark-grown parsley seedlings. Z Naturforsch C 47, 591–600.

Yalpani, N., Raskin, I., 1993. Salicylic acid: A systemic signal in induced plant disease resistance. Trends Microbiol. 1, 88-92.

Yamauchi, K., Yasuda, S., Fukushima, K., 2002. Evidence for the biosynthetic pathway from sinapic acid to syringyl lignin using labeled sinapic acid with stable isotope at both methoxy groups in *Robinia pseudoacacia* and *Nerium indicum*. J Agric Food Chem 50, 3222-3227.

Yamauchi, K., Yasuda, S., Hamada, K., Tsutsumi, Y., Fukushima, K., 2003. Multiform biosynthetic pathway of syringyl lignin in angiosperms. Planta 216(3), 496–501

Yasumatsu, K., 1967. Distribution and bionomics of natural enemies of rice stem borers. Mushi (Suppl.), 33-44.

Yazaki, K., Heide, L., Tabata, M., 1991. Formation of p-hydroxybenzoic acid from p-coumaric acid by cell free extract of Lithospermum erythrorhizon cell cultures. Phytochemistry, 30, 2233-2236.

Ye, Z.H., Kneusel, R.E., Matern, U. Varner, J.E. 1994. An alternative methylation pathway in lignin biosynthesis in *Zinnia*. Plant Cell 6, 1427-1439.

Yin, ZQ., Wong, WS., Ye, WC., Li, N. 2003. Biologically active *cis*-cinnamic acid occurs naturally in *Brassica parachinensis*. Chin Sci Bull 48, 555–558.

Yoshida, K., Kondo, T., Kameda, K., Goto, T., 1990. Structure of anthocyanins isolated from purple leaves of Perilla ocimoides L. var. crispa Benththeir isomerisation by irradiation of light. Agric. Biol. Chem. 54, 1745–1751.

Zamir, LO., Nedea, ME., Gazneau, FX., 1992. Biosynthetic building blocks of *Taxus canadensis* taxanes. Tetrahedron Lett 33, 5235–5236.

Zhang, X-H., Chiang, VL., 1997. Molecular cloning of 4-coumarate:coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood. Plant Physiol 113, 65-74

8. APPENDIX

- **Nucleotide sequence of fragment A:**

```
GCCAAAAGCCCCGACTTAGATCGGTACGACTTGTCATCGATAAGGATGGTGATGTCCGGTGCGGCGCCGATGGGG
AAGGAGCTTGAGGATACAAGTGAGGGCTAAAGTTACCTAAGTGCCAAACTTGGGACAGGGTAAATTCAAATTGAA
TAACGTCTTTAATTGGATTTCTCTGTATATATATACCATATTTCCTAAGTGGTTGAGTGATTCTCATCCTCAATC
ATGTTTTGGTAATTTTAGTCCGGAAAATGCTATAAATACCATATTTGTTGACCAACTTTAAAAAAAAAAAAAAAA
AAAAAAAAAA
```

- **Amino acid sequence and data bank comparison of fragment A:**

```
gi|9651915|gb|AAF91309.1|  4-coumarate:coA ligase 2 [Rubus idaeus]
Length=544
            Identities = (75%)          Frame = +1

Query  1    AKSPDLDRYDLSSIRMVMSGAAPMGKELEDTSEG*SYLSAKLGTG  135
            AK PDL RYDLSSIRMVMSGAAPMGKELEDT       +AKLG G
Sbjct  291  AKCPDLHRYDLSSIRMVMSGAAPMGKELEDTVRA-KLPNAKLGQG  334
```

- **Nucleotide sequence of fragment B:**

```
ATTCTAACACACATCAACTTGGTCACCAGCACGGCCCAGCAGGTGGACCGAGAGAATCCAAACCTATAATTCAAA
GATGACCACGTCGTATTCTACGTGCTGCCGTTGTTTCACATATTGTCGCGCAAAATCGCGCTAAACAGCCTGGTT
GCGAGCATGGGCCGGACTTCTGCTGATGCACAAGTTTAAGATAAATTCTCTTATAGAGCTCATTCTGCAGCTCCT
ACTGCCGGTCCGCAGCGGCGGTGCCACCGCTGGTTATACCGCTGGCTATAAACCCAACGGTGACCCAATTGGTCC
CGAGCTCTATTACGGTGGTGTATTACAGGAGCTGCTCCTCTGAGCAACGACCCTGAGGAAGGACCTAAAAGCCGA
GCTCATCAGGCAGTGTCCCTTAAGCAGTATTGGATCACGGTTACAGGATCACTGCTGCCAATG
```

- **Amino acid sequence and data bank comparison of fragment B:**

```
gi|9651913|gb|AAF91308.1|  4-coumarate:coA ligase 3 [Rubus idaeus]
Length=591
          Identities = (36%)          Frame = +1

Query 1   ILTHINLVTSTAQQVDRENPNL*FKDDHVVFYVLPLFHILSRKIALNSLVASMGRTSADA  180
          ILTH +L+TS AQQVD ENPNL  K D VV  VLPLFHI S    LNS++    R  A
Sbjct 242 ILTHKSLITSVAQQVDGENPNLYLKGDDVVLCVLPLFHIFS----LNSVLLCSLRAGAAV  297

Query 181 QV*DKFS----------YRAHS------AAPTAGPQRRCHRWLYRWL*TQR*PNWSRALL  312
           V  KF           YR         A  +AG +                ++  + +
Sbjct 298 LVMPKFEIGTLLELIQRYRVFCGGGWCLAGDSAGEESMV-------------ADYDLSSI  344

Query 313 RWCITGAAPLSNDPEEGPKSRAHQAVSLKQYWITVTGSLLPM  438
          R  ++GAAPL  + EE  ++R  QAV  + Y +T  G +L M
Sbjct 345 RVVLSGAAPLGKELEEALRNRVPQAVLGQGYGMTEAGPVLSM  386
```

Curriculum vitae

Name: Hussein Mohamed Ramadan

Date of birth: 10.04.1969

Place of birth: Rabta, Libya

Nationality: Libyan

Marital status: Married and two children

School:	1975 – 1981	Elementary School Al-rabta
	1981 – 1984	Junior High School Al-wady Elhay
	1984 – 1987	Senior High School Al-wady Elhay
Studies:	1987 – 1993	Graduate study at Faculty of Pharmacy, Alfateh University
	02.07.1993	Bachelor in Pharmacy, Alfateh University, Tripoli, Libya

Working Experiences:

09.93 – 10.97	Demonstrator in natural products, Alfateh University, and head of medical supply department in National Pharmaceutical Company in Tripoli-Libya.
11.97 – 05.01	Demonstrator in natural Products, Garyounis University, and pharmacist in technical department of I.V infusion factory in Benghazi-Libya
February 2001	Scholarship from Libyan government for Ph.D. work

Language courses:

02.01 – 06.01	German intensive course in Goethe-Institute, Bonn
06.01 – 02.02	English intensive course in Wall-Street-Institute, Bonn and Hannover

PhD. work:

02.02 - 04.06	Institute of Pharmaceutical Biology, Technical University of Braunschweig